

## Themes from Ontology, Mind, and Logic

# Grazer Philosophische Studien

INTERNATIONAL JOURNAL FOR ANALYTIC PHILOSOPHY

*Founded by*

Rudolf Haller

*Edited by*

Johannes L. Brandl (*Universität Salzburg*)

Marian David (*Universität Graz*)

Maria E. Reicher (*Universität Aachen*)

Leopold Stubenberg (*University of Notre Dame*)

*Managing Editor*

Martina Fürst (*Universität Graz*)

*Editorial Board*

Peter Baumann – Monika Betzler – Elke Brendel – Annalisa Coliva

Victor Caston – Thomas Crisp – Katalin Farkas – Dagfinn Føllesdal

Volker Gadenne – Christopher Gauker – Hanjo Glock – Kathrin Glüer-Pagin

Robert M. Harnish † – Katherine Hawley – Christoph Jäger – Tom Kelly

Jaegwon Kim – Peter Koller – Nikola Kompa – Wolfgang Künne

Martin Kusch – Karel Lambert – Sandra Lapointe – Keith Lehrer

Hannes Leitgeb – Joseph Levine – Georg Meggle – Thomas Mormann

Edgar Morscher – Herlinde Pauer-Studer – Christian Piller – Marga Reimer

Tobias Rosefeldt – Edmund Runggaldier – Carolina Satorio – Werner Sauer

Susanna Schellenberg – Hans-Bernhard Schmid – Benjamin Schnieder

Gerhard Schurz – Geo Siegart – Peter Simons – Barry Smith

Thomas Spitzley – Matthias Steup – Mark Textor – Udo Thiel – Thomas Uebel

Maria van der Schaar – Ted Warfield – Charlotte Werndl – Nicholas White

VOLUME 91

The titles published in this series are listed at [brill.com/gps](http://brill.com/gps)

# Themes from Ontology, Mind, and Logic

*Present and Past*

Essays in Honour of Peter Simons

*Edited by*

Sandra Lapointe



BRILL  
RODOPI

LEIDEN | BOSTON

Grazer Philosophische Studien is published with the support of:  
The Faculty of Humanities of the University of Graz and



Library of Congress Control Number: 2015945187

ISSN 0165-9227

ISBN 978-90-04-30224-2 (paperback)

ISBN 978-90-04-30227-3 (e-book)

Copyright 2015 by Koninklijke Brill nv, Leiden, The Netherlands.

Koninklijke Brill nv incorporates the imprints Brill, Brill Hes & De Graaf, Brill Nijhoff, Brill Rodopi and Hotei Publishing.

All rights reserved. No part of this publication may be reproduced, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission from the publisher.

Authorization to photocopy items for internal or personal use is granted by Koninklijke Brill nv provided that the appropriate fees are paid directly to The Copyright Clearance Center, 222 Rosewood Drive, Suite 910, Danvers, MA 01923, USA.

Fees are subject to change.

This book is printed on acid-free paper.

## TABLE OF CONTENTS

Preface . . . . .	vii
-------------------	-----

### PART I: ONTOLOGY

Fabrice CORREIA: Logical Grounding and First-Degree Entailments	3
Ingvar JOHANSSON: Collections as One-and-Many—On the Nature of Numbers . . . . .	17
Kathrin KOSLICKI: In Defense of Substance . . . . .	59
Uriah KRIEGEL: How to Speak of Existence: A Brentanian Approach to (Linguistic and Mental) Ontological Commitment . . . . .	81
Maria Elisabeth REICHER: Computer-generated Music, Authorship, and Work Identity . . . . .	107
Benjamin SCHNIEDER: The Asymmetry of ‘Because’ . . . . .	131
Johanna SEIBT: Non-Transitive Parthood, Leveled Mereology, and the Representation of Emergent Parts of Processes . . . . .	165
Theodore SIDER: Nothing Over and Above . . . . .	191
David WOODRUFF SMITH: On Basic Modes of Being: Metametaphysical Reflections in Light of Whitehead, Husserl, Ingarden, Hintikka . . . . .	217

### Part II: Mind

Kevin MULLIGAN: Annehmen, Phantasieren und Entertaining. Husserl und Meinong . . . . .	245
Mark TEXTOR: Meaning, Entertaining, and Phantasy Judgement . . .	285
Maria VAN DER SCHAAR: The Things We Call True . . . . .	303

PART III: LOGIC

Bob HALE & Crispin WRIGHT: Bolzano's Definition of Analytic Propositions . . . . .	325
Wolfgang KÜNNE: On Having a Property. Corrigenda in Bolzano's <i>Wissenschaftslehre</i> . . . . .	365
Edgar MORSCHER: The Logic of Truth . . . . .	409
Jan WOLEŃSKI: An Analysis of Logical Determinism . . . . .	423

## PREFACE

I spent three years working under Peter's supervision in Leeds between 1997 and 2000. As I entered his office, on the day I handed him my dissertation, he looked at me holding my humble green linen-bound manuscript. That was it. I was done. Without saying a word, he reached for a fresh and crisp copy of his newly arrived paperback edition of *Parts*. He took a pen, scribbled a few words on the title page and handed it to me. It read: "A good book deserves another". To this day I have wondered what other book he was talking about and who might write it.

There is something colossal about Peter's contribution to philosophy, not only on account of its sheer magnitude—he has published over 225 articles on a variety of subjects broad enough to qualify him as an erudite—but also because of its "Stoff". In metaphysics, Peter not only published substantial work on topics including but not limited to: truthmaking, mereology, theories of substance, grounding and existential dependence, patterns, pluralities and tropes; in many cases, much of the discussion today can to a large extent be traced back to his pioneering work. In all his work, Simons develops and defends original views, consistently attempting to articulate a robustly naturalist and nominalist approach. This holds as well for his work on plurals, quantification, meaning and categorial grammar which he often develops in conjunction to historical studies. Indeed, Peter's remarkable acumen as a historian of analytical philosophy has contributed to transforming our understanding of nineteenth and early twentieth century philosophy. He has published extensively on the history of philosophy in Central Europe, offering new and enlightening analyses of Bolzano, Brentano, Husserl, Meinong, Lesniewski and Tarski. He has also written on Frege, the early Russell and the history of logic, mathematics and their philosophy at large. His discussions are formidably clear and accurate, inspiring and humbling. In all this he should be considered as a model for our students and a formidable and precious colleague.

The present volume is meant to pay tribute to Peter's prodigious career on the occasion of his 65<sup>th</sup> birthday. The essays gathered were written by friends, students and other eminent scholars who share the same admiration for his keen intellect and untiring efforts.

Sandra LAPOINTE



PART I  
ONTOLOGY



## LOGICAL GROUNDING AND FIRST-DEGREE ENTAILMENTS

Fabrice CORREIA  
University of Neuchâtel

### *Summary*

I give a characterisation of three relations of logical grounding in sequent format, which I use to highlight some connections between logical grounding and first-degree entailments understood à la Anderson and Belnap.

The view that there is a distinction among inferential connections between those which are explanatory and those which are not has found proponents throughout the history of philosophy—the most well-known being perhaps Aristotle (*Posterior Analytics*, I, 2 and 13) and Bernard Bolzano (*Theory of Science*, esp. II, §198 and IV, §525). In my “Logical Grounds” (Correia 2014) I developed a theory of *logical grounding* which is based on this distinction. For present purposes, grounding may be taken to be a many-one relation between propositions, and the notion may be explicated by saying that some given propositions ground another given proposition when the former propositions’ being true *makes it the case that* the latter proposition is true. A central tenet of “Logical Grounds” is that grounding comes in various kinds, and that one of these kinds is distinctively logical. The paper gives a precise characterisation of logical grounding and shows that the notion can be fruitfully used in certain areas of logical inquiry.

One of the main results put forward in “Logical Grounds” is that various well-known consequence relations—in particular, the relation of classical logical consequence defined on propositional or first-order languages—can be fully characterised in terms of logical grounding. One such relation, associated with the so-called first-degree entailments of Anderson and Belnap (1962, 1963), is actually more closely tied to logical grounding than the other ones. The aim of the present paper is to further the study of these ties. In the first section, I present the theory of logical grounding for propositional languages more or less as it is formulated in “Logical Grounds”. In the second section, I characterise the consequence relation

of interest to us, still relative to propositional languages, and in the final section I present and discuss some important connections between that relation and logical grounding.

### 1. *Logical grounding*

The theory of logical grounding to be formulated here supposes given a standard propositional language with  $\wedge$ ,  $\vee$  and  $\neg$  as primitive connectives. Following common usage, the formulas of the language which are either atoms or negatedatoms are called *literals*, and I use  $\phi$ ,  $\psi$ , etc. for formulas of the language and  $\Delta$ ,  $\Gamma$ , etc. for sets thereof. Standard definitions and notational conventions will be used throughout the paper.

The theory assumes the following basic rules of inference:

$$\begin{array}{lll}
 (\wedge 1) \quad \frac{\phi \quad \psi}{\phi \wedge \psi} & (\wedge 2) \quad \frac{\neg\phi}{\neg(\phi \wedge \psi)} & (\wedge 3) \quad \frac{\neg\psi}{\neg(\phi \wedge \psi)} \\
 (\vee 1) \quad \frac{\neg\phi \quad \neg\psi}{\neg(\phi \vee \psi)} & (\vee 2) \quad \frac{\phi}{\phi \vee \psi} & (\vee 3) \quad \frac{\psi}{\phi \vee \psi} \\
 (\neg) \quad \frac{\phi}{\neg\neg\phi}
 \end{array}$$

Each rule, read from top to bottom, is supposed to encode a link of logical grounding:  $(\wedge 1)$  states that for any  $\phi$  and  $\psi$ ,  $\phi$  and  $\psi$  together ground  $\phi \wedge \psi$ ,  $(\neg)$  states that for any  $\phi$ ,  $\phi$  grounds  $\neg\neg\phi$ , and so on.

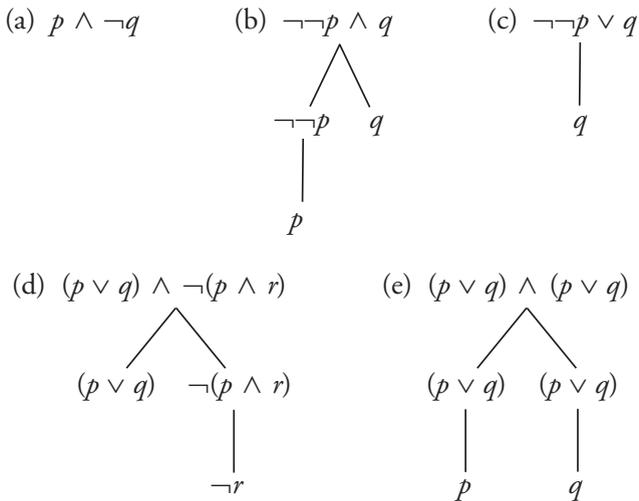
The theory also assumes that these rules are sufficient, in the sense that all the connections of logical grounding between formulas of our language can be described in terms of these rules. There are actually various relations that can be defined in terms of the basic rules, which correspond to various concepts of logical grounding.

Let a *grounding tree* be a rooted tree  $T$  which satisfies the following conditions:

1. Each of  $T$ 's nodes is occupied by a formula;
2. No parent node in  $T$  is occupied by a literal;
3. Given a parent node  $N$  in  $T$ :

- If  $N$  is occupied by  $\phi \wedge \psi$ ,  $N$  has two children, one occupied by  $\phi$  and the other one by  $\psi$
- If  $N$  is occupied by  $\phi \vee \psi$ ,  $N$  has one child, occupied by  $\phi$  or by  $\psi$
- If  $N$  is occupied by  $\neg(\phi \wedge \psi)$ ,  $N$  has one child, occupied by  $\neg\phi$  or by  $\neg\psi$
- If  $N$  is occupied by  $\neg(\phi \vee \psi)$ ,  $N$  has two children, one occupied by  $\neg\phi$  and the other one by  $\neg\psi$
- If  $N$  is occupied by  $\neg\neg\phi$ ,  $N$  has one child, occupied by  $\phi$ .

The following are examples of grounding trees:



A grounding tree *for* a formula is a grounding tree whose root node is occupied by the formula itself, and a grounding tree for a formula is said to be *from* a set of formulas  $\Delta$  iff  $\Delta$  is the set of all the formulas which occupy leaves on the grounding tree. Thus, (a) above is a grounding tree for  $p \wedge \neg q$  from  $\{p \wedge \neg q\}$ , (b) a grounding tree for  $\neg\neg p \wedge q$  from  $\{p, q\}$ , etc. A grounding tree is said to be *degenerate* iff it consists of just one node.

Three concepts of logical grounding may then naturally be defined:<sup>1</sup>

1. Here as in “Logical Grounds”, I treat logical grounding as a relation between sets of formulas and formulas, and I treat it as a non-factive relation (‘ $\Delta$  logically grounds  $\phi$ ’ does not entail ‘ $\phi$  and all the members of  $\Delta$  are true’). Some might demur on both counts. Yet my treatment

**Definition 1.1.** For  $\Delta$  a set of formulas and  $\phi$  a formula:

- $\Delta$  STRICTLY GROUNDS\*  $\phi$ —in symbols:  $\Delta \triangleright^* \phi$ —*iff<sub>def</sub>* there is a non-degenerate grounding tree for  $\phi$  from  $\Delta$ .
- $\Delta$  STRICTLY GROUNDS  $\phi$ —in symbols:  $\Delta \triangleright \phi$ —*iff<sub>def</sub>* there is a covering of  $\Delta$  (i.e. a family of sets whose union is  $\Delta$ ) such that for each  $\Delta'$  in this covering,  $\Delta' \triangleright^* \phi$ .
- $\Delta$  WEAKLY GROUNDS  $\phi$ —in symbols:  $\Delta \trianglerighteq \phi$ —*iff<sub>def</sub>* for some  $\Delta' \subseteq \Delta$ , there is a grounding tree for  $\phi$  from  $\Delta'$ .

One can readily verify that for every set of formulas  $\Delta$  and every formula  $\phi$ :

- $\Delta \trianglerighteq \phi$  iff either  $\phi \in \Delta$ , or for some  $\Delta' \subseteq \Delta$ ,  $\Delta' \triangleright^* \phi$
- $\Delta \trianglerighteq \phi$  iff either  $\phi \in \Delta$ , or for some  $\Delta' \subseteq \Delta$ ,  $\Delta' \triangleright \phi$ .

Thus, the weak relation is definable in terms of either of the strict relations. One can also verify that  $\triangleright^*$  is strictly stronger than  $\triangleright$ , which in turn is strictly stronger than  $\trianglerighteq$ . Further important properties of these three relations are listed below:

Properties of STRICT GROUNDS\*:

1. If  $\Delta \triangleright^* \phi$ , then  $\Delta \neq \emptyset$  and is finite
2. If  $\Delta \triangleright^* \phi$ , then  $\phi$  is not a literal
3. If  $\Delta \triangleright^* \phi$ , then  $\text{Complexity}(\phi) > \text{Complexity}(\psi)$  for any  $\psi \in \Delta$
4. Not:  $\Delta, \phi \triangleright^* \phi$  *Generalised Irreflexivity*
5. If  $\Delta \triangleright^* \psi$  and  $\psi, \Delta' \triangleright^* \phi$  and  $\psi \notin \Delta'$ , then  $\Delta, \Delta' \triangleright^* \phi$  *Restricted Cut*

Properties of STRICT GROUNDS:

1. If  $\Delta \triangleright \phi$ , then  $\Delta \neq \emptyset$  and is finite
2. If  $\Delta \triangleright \phi$ , then  $\phi$  is not a literal
3. If  $\Delta \triangleright \phi$ , then  $\text{Complexity}(\phi) > \text{Complexity}(\psi)$  for any  $\psi \in \Delta$
4. Not:  $\Delta, \phi \triangleright \phi$  *Generalised Irreflexivity*
5. If  $\Delta \triangleright \psi$  and  $\psi, \Delta' \triangleright \phi$ , then  $\Delta, \Delta' \triangleright \phi$  *Cut*
6. If  $\Delta \triangleright \phi$  and  $\Delta' \triangleright \phi$ , then  $\Delta, \Delta' \triangleright \phi$  *Amalgamation*

---

of logical grounding as a relation between sets of formulas and formulas is inessential and could easily be abandoned in favour of other treatments. On factivity, see “Logical Grounds“, 35f.

Properties of WEAK GROUNDING:

1. If  $\Delta \vDash \phi$ , then  $\Delta \neq \emptyset$
2. If  $\Delta \vDash \phi$ , and  $\phi \notin \Delta$ , then  $\phi$  is not a literal
3.  $\phi \vDash \phi$  *Reflexivity*
4. If  $\Delta \vDash \psi$  and  $\psi, \Delta' \vDash \phi$ , then  $\Delta, \Delta' \vDash \phi$  *Cut*
6. If  $\Delta \vDash \phi$ , then  $\Delta, \Delta' \vDash \phi$  *Weakening*

The strict relations satisfy neither Reflexivity nor Weakening, and for that reason they are presumably closer to our intuitive conception of logical grounding than the weak relation. The weak relation is nevertheless theoretically very useful, as we will see below. Whether one of the strict relations is closer than the other one to our intuitive conception of logical grounding I do not know.

Since the basic rules of inference for grounding are all classically valid, the weaker relation is stronger than classical logical consequence: for every set of formulas  $\Delta$  and every formula  $\phi$ , if  $\Delta \vDash \phi$ , then  $\phi$  is a classical consequence of  $\Delta$ . It should actually also be clear that weak grounding is stronger than certain consequence relations which are strictly stronger than classical consequence. One important such relation is the relevant consequence relation associated with Anderson and Belnap's (1962, 1963) first-degree entailments. The connections between weak grounding and that notion of consequence—FDE-consequence, for short—run actually very deep.

## 2. FDE

Let me here characterise FDE-consequence in a general form, and present a proof system which captures the relation so defined. (On the basics of FDE and the related logics mentioned at the end of this section, see e.g. (Priest 2008).)

Let a *valuation* be a distribution of truth-values (T and F) over the atoms of our language. Neither gaps nor gluts are excluded: an atom may be assigned no truth-value at all, or both T and F. Truth ( $\models$ ) and falsity ( $\models$ ) for formulas relative to a valuation  $v$  are then defined recursively as follows:

- For  $\phi$  atomic:  $v \models \phi$  iff  $p$  is assigned T by  $v$
- For  $\phi$  atomic:  $v \models \phi$  iff  $p$  is assigned F by  $v$

- $v \models \phi \wedge \psi$  iff  $v \models \phi$  and  $v \models \psi$
- $v \models \phi \vee \psi$  iff  $v \models \phi$  or  $v \models \psi$
- $v \models \neg\phi$  iff  $v \not\models \phi$
- $v \models \phi$  iff  $v \models \neg\neg\phi$

FDE-consequence is most frequently characterised as a relation between two formulas, and one way of doing it makes use of valuations as just defined and runs as follows:  $\psi \models_{\text{FDE}} \phi$  iff for every valuation  $v$ , if  $v \models \psi$ , then  $v \models \phi$ . I generalise a bit and define it as a relation between two sets of formulas:

**Definition 2.1.** For all sets of formulas  $\Delta$  and  $\Gamma$ :  $\Delta \models_{\text{FDE}} \Gamma$  iff<sub>def</sub> for every valuation  $v$ , if ( $v \models \psi$  for all  $\psi \in \Delta$ ), then ( $v \models \phi$  for some  $\phi \in \Gamma$ ).

FDE-consequence so defined can be proof-theoretically characterised in an elegant way by means of a multiple-conclusion sequent calculus defined by the following axioms and rules:<sup>2</sup>

### System FDE

*Introduction axioms:*

- i1.  $\phi, \psi \vdash \phi \wedge \psi$
- i2.  $\phi \vdash \phi \vee \psi$
- i3.  $\psi \vdash \psi \vee \phi$
- i4.  $\neg\phi, \neg\psi \vdash \neg(\phi \vee \psi)$
- i5.  $\neg\phi \vdash \neg(\phi \wedge \psi)$
- i6.  $\neg\psi \vdash \neg(\phi \wedge \psi)$
- i7.  $\phi \vdash \neg\neg\phi$

*Elimination axioms:*

- e1.  $\phi \wedge \psi \vdash \phi$
- e2.  $\phi \wedge \psi \vdash \psi$
- e3.  $\phi \vee \psi \vdash \phi, \psi$
- e4.  $\neg(\phi \vee \psi) \vdash \neg\phi$
- e5.  $\neg(\phi \vee \psi) \vdash \neg\psi$
- e6.  $\neg(\phi \wedge \psi) \vdash \neg\phi, \neg\psi$
- e7.  $\neg\neg\phi \vdash \phi$

*Structural rules:*

Cut:

$$\frac{\Delta \vdash \Gamma, \phi \quad \phi, \Delta' \vdash \Gamma'}{\Delta, \Delta' \vdash \Gamma, \Gamma'}$$

2. The axioms could, of course, be replaced by rules with zero premisses.

Weakening:

$$\frac{\Delta \vdash \Gamma}{\Delta, \Delta' \vdash \Gamma, \Gamma'}$$

(Notice that the only axioms with a multiple conclusion are the elimination axioms e3 and e6.) The sequents provable in this calculus are said to be *FDE-provable*, and I will write  $\Delta \Vdash_{\text{FDE}} \Gamma$  to say that the sequent  $\Delta \vdash \Gamma$  is FDE-provable. (The same type of notation will be used for other systems below.)

As previously announced:

**Theorem 2.2.** For all  $\Delta$  and  $\Gamma$ :  $\Delta \Vdash_{\text{FDE}} \Gamma$  iff  $\Delta \models_{\text{FDE}} \Gamma$ .

*Proof.* The proof of soundness is straightforward. For completeness, suppose that  $\Delta \not\Vdash_{\text{FDE}} \Gamma$ . Enumerate the formulas:  $\phi_0, \phi_1, \dots$ , and define a series of sets of formulas  $(S_n)_{n \in \mathbb{N}}$  as follows:

- $S_0 = \Delta$ ;
- $S_{n+1} = S_n \cup \{\phi_n\}$  if  $S_n, \phi_n \not\Vdash_{\text{FDE}} \Gamma$ , and  $S_n$  otherwise.

Define  $\Delta^+ \text{ as } \bigcup_{n \in \mathbb{N}} S_n$ . Notice that by construction,  $\Delta^+ \cap \Gamma = \emptyset$ . One can establish that for all formulas  $\phi$  and  $\psi$ :

1.  $\phi \wedge \psi \in \Delta^+$  iff both  $\phi \in \Delta^+$  and  $\psi \in \Delta^+$ ;
2.  $\neg(\phi \wedge \psi) \in \Delta^+$  iff  $\neg\phi \in \Delta^+$  or  $\neg\psi \in \Delta^+$ ;
3.  $\phi \vee \psi \in \Delta^+$  iff  $\phi \in \Delta^+$  or  $\psi \in \Delta^+$ ;
4.  $\neg(\phi \vee \psi) \in \Delta^+$  iff both  $\neg\phi \in \Delta^+$  and  $\neg\psi \in \Delta^+$ ;
5.  $\neg\neg\phi \in \Delta^+$  iff  $\phi \in \Delta^+$ .

Define valuation  $v$  by stipulating that for every atom  $\phi$ ,  $v$  assigns T to  $\phi$  iff  $\phi \in \Delta^+$  and  $v$  assigns F to  $\phi$  iff  $\neg\phi \in \Delta^+$ . Using points 1–5 above, one can prove by induction on the length of the formulas that for all formulas  $\phi$ :  $v \models \phi$  iff  $\phi \in \Delta^+$ , and  $v \not\models \phi$  iff  $\neg\phi \in \Delta^+$ . Since  $\Delta \subseteq \Delta^+$ , it follows that  $v \models \psi$  for all  $\psi \in \Delta$ . On the other hand, since  $\Delta^+ \cap \Gamma = \emptyset$ , it also follows that there is no formula  $\phi \in \Gamma$  such that  $v \models \phi$ . Consequently,  $\Delta \not\models_{\text{FDE}} \Gamma$ .

As an aside, notice that if we alter definition 2.1 by imposing certain conditions on valuations in the definiens, we obtain characterisations of other well-known consequence relations: if one excludes gaps, one gets LP-consequence; if one excludes gluts, one gets K3-consequence; and if one excludes both, one gets classical consequence. In order to obtain an adequate calculus for LP-consequence, it suffices to add the axiom  $\vdash \phi, \neg\phi$  to the calculus for FDE-consequence presented above; for K3 consequence it suffices to add  $\phi, \neg\phi \vdash$  instead; and for classical consequence, it suffices to add both.

### 3. Connections

Let a *situation* be a set of literals. In “Logical Grounds”, I worked with a many-one notion of FDE-consequence and I established the following connection between that notion and weak grounding:

- Given any set of formulas  $\Delta$  and any formula  $\phi$ :  $\Delta \models_{\text{FDE}} \phi$  iff for every situation  $\Lambda$ , if  $(\Lambda \supseteq \psi$  for all  $\psi \in \Delta)$ , then  $\Lambda \supseteq \phi$ .

A more general result concerning many-many FDE-consequence can be established in much the same way:

**Theorem 3.1.** *For all  $\Delta$  and  $\Gamma$ :  $\Delta \models_{\text{FDE}} \Gamma$  iff for every situation  $\Lambda$ , if  $(\Lambda \supseteq \psi$  for all  $\psi \in \Delta)$ , then  $(\Lambda \supseteq \phi$  for some  $\phi \in \Gamma)$ .*

This shows that FDE-consequence can be defined in terms of WEAK GROUNDING, and hence ultimately in terms of either of its strict counterparts. (The same holds of LP-consequence, K3-consequence and classical consequence.)

An almost immediate upshot of this result, which is not stated in “Logical Grounds”, is that:

**Theorem 3.2.** *For every situation  $\Delta$  and every formula  $\phi$ :  $\Delta \supseteq \phi$  iff  $\Delta \models_{\text{FDE}} \phi$ .*

This shows that WEAK GROUNDING can be defined in terms of FDE-consequence when restricted to a relation between sets of literals and arbitrary formulas.

These theorems state deep links between logical grounding as characterised above and FDE-consequence. But one can go significantly fur-

ther thanks to the sequent calculus for FDE introduced in the previous section.

Notice that the seven introduction axioms of that calculus correspond, in an obvious sense, to the seven basic rules for grounding, while none of the elimination axioms corresponds, in the same sense, to a acceptable links of ground. Also, remember that WEAK GROUNDING obeys both a principle of Cut and a principle of Weakening. Would dropping the elimination axioms from the sequent calculus for FDE give us an adequate characterisation of WEAK GROUNDING?

Not quite, for two reasons: first, weak grounding is not many-many, and second, weak grounding is reflexive, and while  $\phi \vdash \phi$  is FDE-provable, its proofs all make use of elimination axioms. Yet consider the system  $\text{FDE}_0$ , defined from system FDE by dropping the elimination axioms and adding an axiom for reflexivity:

### System $\text{FDE}_0$

*Introduction axioms:*

- i1.  $\phi, \psi \vdash \phi \wedge \psi$
- i2.  $\phi \vdash \phi \vee \psi$
- i3.  $\psi \vdash \psi \vee \phi$
- i4.  $\neg\phi, \neg\psi \vdash \neg(\phi \vee \psi)$
- i5.  $\neg\phi \vdash \neg(\phi \wedge \psi)$
- i6.  $\neg\psi \vdash \neg(\phi \wedge \psi)$
- i7.  $\phi \vdash \neg\neg\phi$

*Structural axiom:*

Reflexivity:

$$\phi \vdash \phi$$

*Structural rules:*

Cut:

$$\frac{\Delta \vdash \Gamma, \psi \quad \psi, \Delta' \vdash \Gamma'}{\Delta, \Delta' \vdash \Gamma, \Gamma'}$$

Weakening:

$$\frac{\Delta \vdash \Gamma}{\Delta, \Delta' \vdash \Gamma, \Gamma'}$$

Then  $\models_{\text{FDE}_0}$  is essentially many-one, i.e. for all sets of formulas  $\Delta$  and  $\Gamma$ ,  $\Delta \models_{\text{FDE}_0} \Gamma$  iff  $\Delta \models_{\text{FDE}_0} \phi$  for some  $\phi \in \Gamma$  (the right-to-left direction is immediate thanks to Weakening, and the left-to-right direction can easily be proved by induction on the length of the derivations). And it can be shown that:

**Theorem 3.3.** *For all  $\Delta$  and  $\phi$ :  $\Delta \models \phi$  iff  $\Delta \Vdash_{\text{FDE}_0} \phi$ .*

*Proof.* For the left-to-right direction, by induction on the height of the grounding trees, and for the right-to-left direction, by induction on the length of the derivations.

This being established, it is clear that a more direct characterisation of WEAK GROUNDING is provided by the following calculus for many-one sequents, which differs from the previous calculus only in its structural rules:

### System $\text{FDE}_1$ (for WEAK GROUNDING)

*Introduction axioms:*

- i1.  $\phi, \psi \vdash \phi \wedge \psi$
- i2.  $\phi \vdash \phi \vee \psi$
- i3.  $\psi \vdash \psi \vee \phi$
- i4.  $\neg\phi, \neg\psi \vdash \neg(\phi \vee \psi)$
- i5.  $\neg\phi \vdash \neg(\phi \wedge \psi)$
- i6.  $\neg\psi \vdash \neg(\phi \wedge \psi)$
- i7.  $\phi \vdash \neg\neg\phi$

*Structural axiom:*

Reflexivity:

$$\phi \vdash \phi$$

*Structural rules:*

Cut:

$$\frac{\Delta \vdash \psi \quad \psi, \Delta' \vdash \phi}{\Delta, \Delta' \vdash \phi}$$

Weakening:

$$\frac{\Delta \vdash \phi}{\Delta, \Delta' \vdash \phi}$$

**Theorem 3.4.** For all  $\Delta$  and  $\phi$ :  $\Delta \triangleright \phi$  iff  $\Delta \Vdash_{FDE_1} \phi$ .

*Proof.* Same strategy as for the previous theorem.

The strict relations can also be characterised in a similar way. The case of STRICT GROUNDING is straightforward:

**System FDE<sub>2</sub>** (for STRICT GROUNDING)

*Introduction axioms:*

- i1.  $\phi, \psi \vdash \phi \wedge \psi$
- i2.  $\phi \vdash \phi \vee \psi$
- i3.  $\psi \vdash \psi \vee \phi$
- i4.  $\neg\phi, \neg\psi \vdash \neg(\phi \vee \psi)$
- i5.  $\neg\phi \vdash \neg(\phi \wedge \psi)$
- i6.  $\neg\psi \vdash \neg(\phi \wedge \psi)$
- i7.  $\phi \vdash \neg\neg\phi$

*Structural rules:*

Cut:

$$\frac{\Delta \vdash \psi \quad \psi, \Delta' \vdash \phi}{\Delta, \Delta' \vdash \phi}$$

Amalgamation:

$$\frac{\Delta \vdash \phi \quad \Delta' \vdash \phi}{\Delta, \Delta' \vdash \phi}$$

**Theorem 3.5.** For all  $\Delta$  and  $\phi$ :  $\Delta \triangleright \phi$  iff  $\Delta \Vdash_{FDE_2} \phi$ .

*Proof.* Same strategy again.

The case of STRICT GROUNDING\* is more complicated and leads to a calculus that is less elegant:

**System FDE<sub>3</sub>** (for STRICT GROUNDING\*)

*Introduction axioms:*

- i1.  $\phi, \psi \vdash \phi \wedge \psi$
- i2.  $\phi \vdash \phi \vee \psi$
- i3.  $\psi \vdash \psi \vee \phi$
- i4.  $\neg\phi, \neg\psi \vdash \neg(\phi \vee \psi)$
- i5.  $\neg\phi \vdash \neg(\phi \wedge \psi)$
- i6.  $\neg\psi \vdash \neg(\phi \wedge \psi)$
- i7.  $\phi \vdash \neg\neg\phi$

*Introduction rules:*

For conjunction:

$$\frac{\Delta \vdash \phi \quad \Delta' \vdash \psi}{\Delta, \Delta' \vdash \phi \wedge \psi}$$

$$\frac{\Delta \vdash \phi}{\Delta, \phi \vdash \phi \wedge \phi}$$

For disjunction:

$$\frac{\Delta \vdash \neg\phi \quad \Delta' \vdash \neg\psi}{\Delta, \Delta' \vdash \neg(\phi \vee \psi)}$$

$$\frac{\Delta \vdash \neg\phi}{\Delta, \neg\phi \vdash \neg(\phi \vee \phi)}$$

*Structural rules:*

Restricted Cut ( $\psi \notin \Delta'$ ):

$$\frac{\Delta \vdash \psi \quad \psi, \Delta' \vdash \phi}{\Delta, \Delta' \vdash \phi}$$

**Theorem 3.6.** For all  $\Delta$  and  $\phi$ :  $\Delta \triangleright^* \phi$  iff  $\Delta \Vdash_{FDE_3} \phi$ .

*Proof.* Same strategy again.

Theorems 3.1 to 3.6 constitute substantial elements of our understanding of the relationships between logical grounding and first-degree entailments. I surmise that they are far from telling the whole story, and in particular I would not be surprised if the three relations of logical grounding involved in these theorems could be neatly characterised in terms of FDE-consequence.<sup>3</sup>

---

3. Material from which this paper stemmed has been presented at the *Kit Fine Conference* (Varano Borghi, Italy, July-August 2013), the workshop *Groundedness in Semantics and Beyond* (Oslo, August 2013) and the workshop *Proofs That and Proofs Why* (IHPST, Paris, November 2013). I am grateful to the audiences of these events for helpful comments and discussions. This work was carried out while I was in charge of the Swiss National Science Foundation projects ‘Grounding—Metaphysics, Science, and Logic’ (Neuchâtel, CRSII1-147685) and ‘The Nature of Existence: Neglected Questions at the Foundations of Ontology’ (Neuchâtel, 100012-150289), and a member of the Spanish Ministry of Economy and Competitiveness ‘The Makings of Truth: Nature, Extent, and Applications of Truthmaking’ (Barcelona, FFI2012-35026).

## REFERENCES

- Anderson, Alan Ross and Belnap, Nuel D. 1962: "Tautological entailments". *Philosophical Studies* 13, 9–24.
- Anderson, Alan Ross and Belnap, Nuel D. 1963: "First-Degree Entailments". *Math. Annalen* 149, 302–319.
- Aristotle 1994: *Posterior Analytics*, 2nd edition. Oxford: Oxford University Press.
- Bolzano, Bernard 2014. *Theory of Science*. Oxford: Oxford University Press.
- Correia, Fabrice 2014: "Logical Grounds". *The Review of Symbolic Logic* 7(1), 31–59.
- Priest, Graham 2008: *An Introduction to Non-Classical Logic: From If to Is*. Cambridge: Cambridge University Press.



COLLECTIONS AS ONE-AND-MANY  
ON THE NATURE OF NUMBERS

Ingvar JOHANSSON  
University of Umeå

*Summary*

Peter Simons has been rather alone in the modern philosophy of mathematics to argue that the natural numbers should be regarded as *properties of* multitudes or collections. This paper, however, sides with Simons, but it modifies his property view by adding the notion of imposed collection boundaries and accepting fictional collections. Although partly inspired by the Husserl of *Philosophy of Arithmetic*, Simons dismisses Husserl's talk of psychological acts of collective combination, but this paper saves them by dressing them in modern cognitive clothes. Hereby, a reasonable *partially* constructivist notion of the natural numbers emerges.

The most naive opinion is that according to which a number is something like a heap, a swarm in which the things are contained lock, stock and barrel. Next comes the conception of number as a property of a heap, aggregate, or whatever else one might call it.

Frege, Review of Dr. E. Husserl's *Philosophy of Arithmetic* (Frege 1972 [1894], 323)

In the early 1980s, Peter Simons interrupted my innocent slumber about the notion of set and the possible views of numbers that I had been taught as a philosophy student in the mid-1960s. When reading his papers “Number and Manifolds” and “Plural Reference and Set Theory” (Simons 1982a, 1982b) something entered my mind that I have thought about now and then ever since. Also, I have with interest and sympathy followed how he has continued to work on these his early ideas. Now I have systematized my related reflections. I end up on his side against all the usual analyses of numbers, but nonetheless there is a gap between our views. Of course I find it unfortunate that his early-Husserl and early-Russell inspired view of

numbers has not yet found a reasonable place in the discussions of mainstream philosophy of mathematics (Husserl 2003 [1891], Russell 2006 [1903]). Apart from the last section, I will discuss only the natural numbers, but mostly for simplicity's sake use the unqualified notion 'number'.

### 1. *Peter Simons on numbers*

According to Euclid's once famous definition in *Elements Book VII*, a number is a multitude composed of units (Euclid 300 BC). Peter Simons' standing basic anti-Euclidean, anti-Platonist, and anti-Fregean idea about numbers is this:

I maintain that [...] number is a property of external things of a kind which I shall call *manifolds* [also: multitudes]. In this I shall basically agree with Husserl against Frege, but the theory involved will perforce take account of Frege's objections to such a theory. (Simons 1982a, 161)

Simons is both a naturalist (meaning: everything that exists exists in our spatiotemporal world) and a nominalist (meaning: there are only spatiotemporally localized particulars; there are no non-localized universals, propositions, sets, or numbers). I think the driving force behind his philosophy of numbers comes out well in this quotation:

One of the biggest problems facing a naturalist is to account for mathematics, a vast, ancient and noble discipline with millions of practitioners and thousands of books full of results, many of them of great subtlety and beauty. (Simons 2011, 1)

A manifold in the quoted sense is something completely stripped of all relations between the objects that make it up; if there are pure relations as objects in the manifold, they are kept distinct from their relata. The very term 'manifold' used by Simons above is only one among several possible; some others are 'multitude', 'collection', 'class', 'aggregate', and 'plurality'. In his contribution to the Michael Dummett volume of the Library of Living Philosophers, "What Numbers Really Are" (Simons 2007), and later, Simons has instead of 'manifold' chosen to use the term 'multitude'. Nowadays he presents his old original idea by saying:

*My view is that numbers are certain non-distributive formal properties of multitudes.* (Simons 2007, 233; hyphen added)

Being non-distributive means not being a property of each of the members of the multitude separately, but only of all of them collectively. Being a *formal* property does here not mean being a formal-logical property; it means being a property that can meaningfully be ascribed to radically different kinds of ontological entities. A multitude of pebbles (physical entities) can have the property of being seven, a multitude of toothaches (mental entities) can have this property, a multitude of propositions (abstract entities) can be seven, and so can a multitude of novel figures (fictional entities), but of these kinds of entity only a multitude of physical entities can have the non-formal properties of weight and volume.

Simons' multitudes are in several respects similar to sets and mereological sums, respectively, but it is the differences that now are of interest. Multitudes differ from both sets and mereological sums in being nothing over and above its members. If the members are concrete, so is the multitude. A set, on the other hand, is always an abstract object distinct from its members, and a mereological sum is always a mereological individual distinct from its parts (Simons 2011, 5f.). Furthermore, there is always a null set, and in some mereological systems a "null item," but there can never be an empty multitude. As Simons says: "I claim that you cannot accept the existence of each of the three individuals A, B and C but deny existence of the trio ABC. In this multitudes differ from sets and also from mereological sums" (Simons 2011, 15).

I think the differences can summarily be put like this: *a set is more than the sum of its members/elements, a mereological fusion is the sum of its members/parts, and a multitude is simply its members.*

The early Peter Simons says, as noted, that he "basically agree[s] with Husserl against Frege," but the phrase 'against Frege' is important. He does not in general agree with Husserl; to the contrary:

[M]y chief disagreement with Husserl is over his contention that pluralities [multitudes] are *constituted* as such by acts of collective combination, and accordingly are higher-order, *categorical* objects. (Simons 1982a, 162)

From the start, Simons attempts to create an axiom system that can underpin his view of what numbers are. At first he calls it 'manifold theory' (Simons 1982b, 220), but later of course 'multitude theory' (Simons 2011, 14). However, there is here more than a change of name. For a long time he aligns with the logician S. Leśniewski, because he regards Leśniewski as (in contradistinction to Frege) accepting plural terms in his

logic. At least at first sight, plural subject terms such as ‘we’, ‘they’, ‘these Xs’, ‘the Xs’, ‘*a*, *b*, and *c*’, etc. refer to multitudes.

The early Simons says: “The existing system of logic which our system most nearly resembles is Leśniewski’s Ontology, sometimes called the calculus of names” (1982b, 238). Today, however, Simons has said farewell to Leśniewski (Simons 2011); the story of how he met him is told in (Simons 2013b, 227f.). The crucial point is that Leśniewski’s calculus allows only first-order multitudes, but today Simons thinks that some problems in the philosophy of mathematics cannot be solved without the introduction of multitudes of higher order. He argues that such multitudes allow his multitude theory “to provide models in a way similar to that of set theory” (Simons 2011, 2), and so ground a logical notion of *following from*. He claims, though, that his staunch nominalism remains the same:

[U]nlike sets, multitudes, *of whatever order* [italics added], are concrete if their members are concrete, and are necessarily such that if their members exist, then they exist. So multitudes are nominalistically acceptable, and if all their urelements (ultimate individual members) are natural, so are they. This closes the gap in the naturalistic account of mathematics and allows a naturalist to be a formalist with a clear ontological conscience. (Simons 2011, 2f.)

Peter Simons and I have got along quite well, but never, as he once put it, “see[n] eye to eye on all topics metaphysical” (Simons 2013a, 517). Like him I am a naturalist, but unlike him I am not a nominalist. I am a realist, although not a Platonist; I think the universals I believe in exist in our spatiotemporal world and nowhere else. However, in contradistinction to the best known contemporary immanent realist, D.M. Armstrong (Armstrong 1978), I think there are *both* universals and tropes; tropes then being spatiotemporally localized instances of universals. This means that I am not as ontologically far away from Simons’ special “nuclear” trope nominalism (Simons 1994) as an unqualified nominalist-realist labeling may indicate.

In the philosophy of numbers, I have because of our common naturalism always felt as being close to his views, and below comes my attempt to sort out where I agree and disagree with him. Let it be added that the views I will put forward are somewhat tentative, and that Simons, with his new acceptance of higher-order multitudes, seems at the moment to waver a little, too. I will not discuss his axiomatizations; only informally discuss the very concept of multitude (manifold, collection, class, aggregate, plurality) and its possible referents.

I completely agree with Simons' view that natural numbers should be regarded as being first-order *properties of something*. That is, not as being some kind of property bearing thing-like *objects* such as Platonic numbers, spatiotemporal numerals, extensions of concepts (Frege-unsaturated or not), and sets of sets. Furthermore, the properties in question are neither *structures* in the sense of modern mathematical structuralism (Bostock 2009, ch. 6.4; Horsten 2012, sect. 4), nor some sort of *relations* (Bigelow 1988, Michell 1993, Armstrong 2010, ch. 13), and nor *kinds* (Lowe 2006, 81–83). On the view to be defended, the natural numbers are in a curious sense monadic properties, although of course (number 1 aside) not of single things, but of multitudes. Let's call it *the property view of the natural numbers*.<sup>1</sup>

This view does of course allow talk of numbers in abstraction from the relevant property bearers, the multitudes. It is no more odd to say that 7 lies in-between the numbers 6 and 8, without mentioning a property bearing multitude, then it is to say that typical red lies in-between the colors light red and dark red without mentioning a property bearer such as a color spot. Furthermore, as soon as operations on numbers are allowed, then the numbers may themselves in turn be ascribed (operation-dependent) properties, as in the sentence '5 is a prime number'. The property view of numbers does not cancel the distinction between pure and applied arithmetic, a fact that will come out more clearly later (Section 6).

With respect to how I differ from Simons, I have already said that I have a realist and Simons a nominalist view of properties, but I also differ from him in being sure that one can make semantic-philosophical sense of talk of fictional objects.<sup>2</sup> More precisely, I think that one can allow also multitudes of fictional objects as property bearers for numbers (Section 7). However, the most important difference between us in relation to the philosophy of number is that whereas Simons completely dismisses Husserl's collective combinations, I will save them by dressing them in modern cognitive clothes.

---

1. This view is also argued for by (Yi 1998). He presents it by criticizing the relational account in (Bigelow 1988). He is aware of Simons as a forerunner, but in a long footnote (25) he complains about ambiguities in (Simons 1982a). I find him over-complaining. About Yi, see also footnote 12 below.

2. My trope accepting realism is best defended in (Johansson 2014a) and my view of fictions in (Johansson 2010).

## 2. *The naturalist Husserl on numbers*

In what follows, partly out of respect for Peter Simons, I will not use his preferred term, ‘multitude’, and thereby try to add new connotations to it. Instead I will talk about *collections*. My views come much closer than Simons’ to those put forward by Edmund Husserl in *Philosophy of Arithmetic* (Husserl 2003 [1891]); a book that in most of the twentieth century has suffered from a serious misinterpretation (see quotation below). I guess it is because of this fact that the book is not even mentioned in ordinary introductions to the philosophy of mathematics; see e.g. (Körner 1960, Bostock 2009, Horsten 2012).<sup>3</sup> In his recent book about Husserl’s philosophical development, the Husserl expert J.N. Mohanty says:

Ever since Frege’s well-known review of the *Philosophie der Arithmetik*, it has been usual to ascribe to Husserl’s work a wrong-headed psychologistic position, which—so the story goes—Husserl later on, partly under the influence of Frege’s criticism, retracted. Recent researchers have shown that this story is mistaken. The thesis of *Philosophie der Arithmetik* is not “psychologistic” in the pejorative sense; it did not *reduce* numbers, for example, to subjective processes. [...] Yet if the thesis is not psychologistic, it is not also what could be, in contrast, called purely logical. (Mohanty 2008, 4)

*Philosophie der Arithmetik* was not translated into English until this century, and the translator comments:

We can see, then, that Husserl made a disastrous choice of terminology in deciding to call the collective combination a “psychical” relation and describing it as having a “psychological nature.” [—] Husserl was, in fact, *never* guilty of “Psychologism” with respect to numbers and their laws, nor in any sense in which he later rejected and refuted it. (Willard 2003, xxi and xxvii)

Readers who want an analytic-philosophically related introduction to Husserl’s *Philosophy of Arithmetic* should read (Bell 1990, 31–84); Bell is of the same opinion as Mohanty and Willard.<sup>4</sup> The subtitle of Husserl’s book is “psychological and logical investigations” (where ‘logic’ is understood in such a broad sense that it encompasses investigations of most kinds

---

3. Of course, noted by Simons: “The history of the philosophy of mathematics during the golden years of 1879–1939 hardly ever mentions Husserl” (Simons 2010, v).

4. Of the opposite view is a recent book on Husserl’s early philosophy of mathematics, (Centrone 2010, xii). The author, however, says she “will approach Husserl’s first book from a perspective that is orthogonal to the psychologism issue” (ibid.). Unfortunately, this also makes the book orthogonal to this paper.

of abstract entities). In order to understand arithmetic epistemologically and ontologically, the early Husserl thinks that two combined kinds of investigations are needed. With this I agree. I will, though, exchange his views about psychological acts with views about cognitive activity in general. I will within block parentheses insert such a change directly in quotations, too.

### 3. *Kripke's lasso catches and Husserl's collective combinations*

Peter Simons takes the sets of set theory, whatever nature the elements have, to be *abstract* objects that do not coincide with their elements, whereas he regards multitudes to be nothing but all the elements, whatever nature they have. Here comes Cantor's famous characterization of sets (quoted from Simons):<sup>5</sup>

By a 'set' we understand any collection into a whole  $M$  of definite and well distinguished objects  $m$  of our intuition or our thought, which are called the 'elements' of  $M$ .

Unter eine "Menge" verstehen wir jede Zusammenfassung  $M$  von bestimmten wohlunterschiedenen Objekten  $m$  unserer Anschauung oder unseres Denken (welche die "Elemente" von  $M$  genannt werden) zu einem Ganzen. (Simons 2005, 143)

According to this characterization, a set is a constructed *whole* ("einem Ganzen"); it is the product of a collecting ("Zusammenfassung"). But after the collecting, the set is regarded as something in itself, moreover as an abstract object. In contemporary discussions about the nature of sets, the phrase 'Kripke's lasso' is sometimes used, e.g. by (Forster 2008). In such cases, it is also in contemporary set theory openly spoken of collecting processes. George Boolos says:

For when one is told that a set is a collection into a whole of definite elements of our thought, one thinks: Here are some things. Now we bind them up into a whole.\* Now we have a set. We don't suppose that what we come up with after combining some elements into a whole could have been one

---

5. Already in *Philosophy of Arithmetic* Husserl calls Cantor a "mathematical genius" (Husserl 2003, 121n3). They were colleagues and close friends in Halle in the late 1880s and in the 1890s. Cantor was a member of the dissertation committee for Husserl's Habilitation thesis, which, later in revised form, became the first four chapters of *Philosophy of Arithmetic*.

of the very things we combined (not, at least, if we are combining two or more elements).

\* We put a “lasso” around them, in a figure of Kripke’s. (Boolos 1971, 220)

When the lasso metaphor is accepted, the lasso must be said to fulfill two functions in relation to its catches. It functions as a boundary of the caught set, and it functions as something that turns the set into an abstract entity. I think it is good to have this contemporary constructivist or semi-constructivist view of sets in the back of one’s mind when, now, I turn to the early Husserl’s constructivist notion ‘collective combination’ (“kollektive Verbindung”), which, as said, Simons wholly dismisses.

Husserl takes great pains to show that there is a way of combining (“verbinden”) objects into a unity, a totality (collection), which does not in any way rely on *any* kind of relations between the objects combined.<sup>6</sup> He summarizes: “When we think particular contents ‘together’ in the manner of a totality, this ‘together’ does not permit itself to be resolved into any other relations, so as to be defined by them” (Husserl 2003, 69). Sometimes, instead of ‘combine’ Husserl uses ‘colligate’ (“kolligieren”), a word that has almost left contemporary vocabulary. For Husserl, as Frege says in his review, “Collective connection [combination] is a relation *sui generis*” (Frege 1972, 322), and the relation is mind-dependent.

Next I will highlight Husserl’s expression ‘unitary representation of their totality’:

It was clear to begin with that the specific nature of the particular objects which are gathered in the form of a multiplicity [“Vielheit”] could contribute nothing to the content of the respective general concept. The only thing that could come into consideration in the formation of these concepts was the *combination* of the objects in the *unitary representation of their totality* [italics added; “einheitlichen Vorstellung ihres Inbegriffes”].<sup>7</sup> (Husserl 2003, 67)

Where the set theorists Cantor and Boolos talk about “collections into *wholes*,” the philosopher Husserl talks about combinations into *unitary*

6. In (Centrone 2010), see note 4, Husserl’s “kollektive Verbindung” is translated as ‘collective *connection*’ instead of ‘collective combination’. I don’t think this is good, since ‘connection’ easily gives rise to the false view that there are pre-given or created relations *between* the objects of a collection.

7. The relation between ‘totality’ and the German ‘Inbegriff’ is at another place in a footnote commented on by the translator: “The point, of course, does not come through in English. ‘Inbegriff,’ which we translate as ‘totality,’ involves *Begriff* or ‘concept,’ which in turn involves *greifen* or ‘to grasp’”; Willard in (Husserl 2003, 100n2). That is, the constructivist element that Husserl builds into his very notion ‘Inbegriff’ is lost in the translation into ‘totality’.

*representations*, i.e., about unitary representations of collections. Here are two other typical quotations (and now starts my insertions of ‘cognitive activity’):

In such cases the contents are just simply thought [cognized] “together,” i.e., as a totality. But in no wise are they truly disjoined or unrelated. To the contrary, they are *joined by means of the psychical act [cognitive activity] holding them together*. It is only that *within* the content of that act all perceptible unification is lacking [first italics added]. (Husserl 2003, 76)

So testimony from many sources—and, above all, from inner experience [cognitive experience] itself—tells us that we must decide in favor of the second view mentioned, according to which collective unification [“Einigung”] is not intuitively given *in* the representation content, but instead has its subsistence only in certain *psychical acts [cognitive activity] that embrace the contents in a unifying manner* [the second italics added; “welche die Inhalte einigend umschliessen”]. (Husserl 2003, 77)

The terms ‘psychical act’ and ‘inner experience’ as here used are essential to the charge that in *Philosophy of Arithmetic* Husserl is defending a reductively psychologistic view of numbers. However, what it amounts to is only psychologism (cognitivism) with respect to the unification of non-connected objects or contents. After the unification, it is by no means for the unifier to decide how many objects the totality (the collection) contains.

The activity of collective combining is only one of two kinds of activities that Husserl finds necessary for the apprehension and constitution of numbers. The other kind of psychic (cognitive) activity is a certain kind of abstraction, one in which abstraction should not be understood as any literal taking away of contents. He specifies: “To disregard or abstract from something is merely to give it no special notice” (Husserl 2003, 83). For instance, if we have cognitively unified a chair, a flower, and a stone into an apprehended collection of three concrete things, we can then stop to give “special notice” to the features that ground their different classificatory concept labels, and simply regard them as three “somethings” (the scare quotes are Husserl’s).<sup>8</sup> Husserl is using “something” (“etwas”) in a sense

---

8. Frege makes in his review the following remark about Husserl’s contentless “something”: “From each object we finally derive something which is completely without restrictions on its content; but the something derived from the one object nevertheless does differ from that derived from the other object, although it is not easy to say how” (Frege 1972, 324). It seems to me as if Frege has not really taken seriously how Husserl wants his abstraction procedure to be under-

that corresponds well to the contemporary philosophical term ‘entity’, but I will in this section continue to use his “something.” Since this term can have anything whatsoever as its referent, it lacks ordinary differentiating conceptual content; and this is one of the reasons behind Husserl’s rejection of Frege’s view that numerical predications about things in the world are assertions about concepts.

Husserl’s view of the natural numbers that he regards as perceivable (i.e. those not larger than twelve (Husserl 2003, 202)) looks like this:

Let us look once more, then, at the psychological [cognitive] foundation of the number concepts.

*According to our view two things make up the concept of number: 1) the concept of “collective unification [combination]” and 2) the concept of “something.”* (Husserl 2003, 355)

Husserl does not regard the mentioned psychological investigations as being purely scientific-psychological. They are at the same time also both epistemological and ontological; let me explain the epistemological aspect first.

At the time, Husserl is an empiricist, although of an almost forgotten (Brentanian) breed. He thinks like Humean empiricists that all meaningful concepts must have some foundation or grounding in the perceptually given, but unlike them he is not a nominalist. He thinks that there are both universals (called ‘abstract concepts’ and ‘abstracta’) and instances of universals (sometimes called ‘Moments’ and sometimes ‘conceptual objects’).<sup>9</sup> I think an analogy between Hume and Husserl can make clear what Husserl’s empiricism in relation to numbers amounts to.

According to Hume, all truly meaningful ideas are at bottom founded on simple impressions. This fact, however, does not for him rule out the possibility of meaningful representations of non-existing entities, i.e., of fictions such as “winged horses, fiery dragons, and monstrous giants” (*Treatise* 1.1.3.4). His view requires only that such complex ideas can be resolved into simple ideas that once have had simple impressions as their referents. Moreover, he even allows as meaningful certain simple ideas that have not been preceded by any corresponding simple impression,

---

stood; in fact there are “restrictions on content” that make the objects differ, but they are not attended to. To Husserl, this notion of “something” seems to have been the point of departure for a lifelong grappling with what it means to be a formal concept; about this see (Mulligan 2013).

9. In *Logical Investigations* (Husserl 1970) Husserl uses other names for universals: ‘Species’, ‘Idea’, and ‘universal object’ (Johansson 2010).

if only their possible referent is meant to be closely similar to some earlier simple impressions (among Hume scholars, the famous case of “the missing shade of blue”). Hume posits a faculty of imagination that he ascribes the following abilities: to create ideas that are copies of impressions, to create ideas of the missing-shade-of-blue kind, to unite impressions and ideas into larger complexes, and to separate such complexes into their parts.<sup>10</sup>

In analogy with Hume’s distinction between ideas preceded and not preceded by directly corresponding impressions, Husserl distinguishes between two kinds of number concepts, *authentic* and *symbolic* ones; the latter have to be grounded in the former in order to be meaningful representations. Authentic number concepts (the natural numbers  $\leq 12$ ) can as referents have phenomena that can be found in perceptual experiences when the two operations of collective combination and abstraction into “somethings” have modified a given perception. Symbolic number concepts, on the other hand, are created by arithmetic operations on the already given authentic number concepts, and can *only* be “representations by means of signs” (Husserl 2003, 205). Based on this distinction, Husserl divides *Philosophy of Arithmetic* into two parts; the first is called “The authentic concepts of multiplicity, unit and whole number,” and the second “The symbolic number concepts and the logical source of cardinal arithmetic,” respectively. Just like Hume, Husserl takes a faculty of imagination for granted.

What then is Husserl’s ontological stance with respect to the referents of the authentic number concepts? As far as I can see, the most reasonable interpretation is this. Both the mind-independent world and the primordial perceptual field contain multiplicities of various kinds, but they do not in themselves contain any numbers. The existence of such multiplicities, however, is a necessary condition for the existence of numbers, but there are two other necessary conditions, too. There must also be results of the operations of collective combination and abstraction (the “somethings”). Jointly, these three necessary conditions (multiplicity, collective combination, “somethings”) make up a sufficient condition for an authentic natural number concept to be applicable. Husserl’s view implies that the referents of authentic number concepts contain one objective part, that of a given multiplicity, and two created subjective-psychological parts (purely cognitive parts).

---

10. The details of my interpretation of Hume can be found in (Johansson 2012).

Equally important to Husserl's analysis of numbers is a distinction between *abstract* and *general* concepts, which he accuses Frege for overlooking (Husserl 2003, 175). I understand him as identifying abstract concepts with universals and general concepts with general terms nominalistically understood, i.e. as mere names for multiplicities of particulars; the general concepts are also called 'general *names*'. He never himself pauses to explain the distinction, but he is very explicitly claiming that the referents of number concepts are not abstract concepts (universals). In a section called "The numbers in arithmetic are not abstracta" he states:

Thus, 5 does not signify the concept (the abstractum) *five*; but rather 5 is a general name (or else a calculational sign) for any arbitrary group as one falling under the [general] concept *five*. (Husserl 2003, 191).

And in the preceding chapter he has said:

Considered formally, number and concrete group are related as are [general] concept and conceptual object [instance of universal]. *Thus number relates, not to the [abstract] concept (Begriff) of the enumerated objects, but rather to their totality (Inbegriff)*. (Husserl 2003, 174)

That is, the authentic (general) number concept 5 does neither in pure nor in applied arithmetic refer to a universal, but always to a totality of "somethings." One of his reasons is that if a number concept had a number universal as its referent, then, since there is by definition only one universal of each kind, an addition such as  $5 + 5$  would make no sense (ibid., 174f.). This observation functions as a *reductio ad absurdum* of all kinds of number Platonism.<sup>11</sup>

If one already has recourse to the natural number concepts, then one can of course say that a given concrete multiplicity must have as many members as it, after the abstraction procedure, has "somethings." If we have 5 pebbles, we have also 5 "somethings"; and vice versa, if we have 5 "somethings" that happen to be pebbles, we also have 5 pebbles. Husserl's point is that the number concept *directly* connects only to the collective combination of "somethings," and only indirectly via the "somethings"

---

11. An analogous argument, it might be noted, has been given for the view that so-called *structural universals* cannot be analyzed as structures of universals, but requires the introduction of tropes (Campbell 1990, ch. 2.7). To be a structural universal in the sense at issue is to be a property that can be characterized by the form 'having  $n$  constituents of kind  $K$ ' ( $n > 1$ ). To be a water molecule, for instance, is to have the property of having 2 hydrogen atoms as constitutive parts, i.e., we have  $1 + 1$  hydrogen atoms. I find both Husserl's and Campbell's arguments valid.

to the collective combination of the pebbles. About this fact, he is very outspoken (Husserl 2003, 174). His view can in the linguistic mode of presentation be put like this: *The sentence 'There are 5 pebbles' is shorthand for 'There is a collective combination of 5 "somethings," and all these "somethings" are pebbles'.*

Ontologically, authentic numbers, i.e., referents of an authentic number concept, can be found in perceptions modified by the operations of collective combination and abstraction. Husserl's view of *symbolic* number concepts can now, taking the Hume analogy for granted, easily be stated. He finds, like Hume, no general semantic problem with concepts that may have merely fictional referents, if only in some way or other the words and symbols used can be linked back to some perceptual phenomena. Therefore, ontologically, symbolic number concepts can refer both to collective combinations of spatiotemporally existing "somethings" and to collective combinations of fictional "somethings."

Simons claims, I repeat, that natural numbers are *formal properties* of multitudes. Husserl comes in my explication close to saying that, first, natural numbers are *properties* of collective combinations; and since the relevant such combinations are combinations of "somethings," he also, second, comes close to saying that numbers are *formal* features of collective combinations of concrete entities. However, he never says either of these things, not even implicitly. Therefore, there is a notable difference between Simons and Husserl. Simons posits *concrete multitudes that lack a collective combination*, and he claims that they have a number *as a property*; Husserl posits *concrete multitudes partly constituted by a collective combination*, which as such he ascribes numbers, not saying whether the numbers should be regarded as properties or as something else. I think there is a position possible in-between Simons' and Husserl's; and a philosophically more reasonable one for that.

I think that Husserl in his way, and Simons in his, has overlooked two connected facts. The first fact is that Husserl's collective combinations and Simons' multitudes (what I call 'collections') must on closer analysis be regarded as simultaneously being both one and many, or being one-and-many. Second, in order to see this, one has to note that collective combinations, multitudes, and collections need an imposed boundary in order to come into existence. Husserl explicitly regards collective combinations as partly psychologically [cognitively] constituted, and Simons implicitly regards multitudes as self-constituting ("they require nothing more for their existence than the existence of their members" (Simons 2011, 17)),

but there is a more credible third position. It will be presented in the next two sections.

#### 4. *Collections as one-and-many*

The heading above alludes to an old subsection heading of Peter Simons: “Classes as many and as one” (Simons 1982b, 200). The latter distinction is introduced by Russell in *Principles of Mathematics*:

In the preceding chapter we regarded the class [collection] as formed by all the terms [members], but usage seems to show no reason why the class [collection] should not equally be regarded as the whole composed of all the terms [members] in those cases where there is such a whole. The first is the class [collection] as many, the second the class [collection] as one. (Russell 2006 [1903], 139)

The “usage” Russell is referring to, he has exemplified with words such as ‘the army’ (compare ‘the soldiers’), ‘the navy’ (compare ‘the sailors’), and the Cabinet (compare ‘the Cabinet Ministers’) (Russell 2006, 68). Read in isolation, the quotation above may invite the view that classes [collections] are entities that, necessarily, have two aspects, a unitary aspect (the class as one) and a plural aspect (the class as many). This comes close to the kind of view of collections that I soon shall try to make sense of, but it is neither Russell’s nor Simons’. To both, a class as many and a class as one are simply two distinct objects, not two necessary conditions for a collection [class, multitude] to be able to exist. I endorse the following view of Husserl: “It is a fact that we are often in a position to conceive of one and the same object as one and as many, as we wish” (Husserl 2003, 163).

According to Simons (leaving from now on Russell aside), classes as one are the abstract objects that set theory takes care of, and classes as many are the multitudes he wants to highlight and create an axiomatization of. I find Simons’ view of the distinction unsatisfactory, and I will use the rest of this section to explain in what way I find it so.

Trivially, where there are several collections, the number of them can in principle be counted. And since they can be counted, there must be some sense in which they are not only collections as many, but also collections as one. It is only *as one* that each collection is counted. Therefore, I shall first try to make clear what I want the term ‘collection as one’ to designate, and then judge whether collections as one can be neglected the way Simons neglects them.

Many collections allow for a distinction between qualitative and numerical identity. When a collection is defined partly or wholly by an intension (i.e., defined as a collection of Xs, where X is a concept or description) this is always the case. For example, if someone talks about a certain collection of five pebbles, then the mere way of speaking indicates that there can be other individual collections falling under the concept 'five pebbles'. If two collections of five pebbles are picked out, then there are two numerically different but qualitatively identical collections; or, in other words, two collections of the same kind.

On the other hand, when a collection is defined by an enumeration, there is seldom need to talk about anything else than the numerical identity of it. Think, for example, of a collection that a speaker picks out by the expression 'this chair, this flower, and this stone'. Even if the collection conforms to the description 'a chair, a flower, and a stone', the speaker may have meant the collection of the three entities pointed at independently of whether they have been correctly classified or not.

In what follows, if nothing to the contrary is said, when talking about the identity of a collection, I am talking about its numerical identity, even though of course its qualitative identity may be of relevance for the numerical identity. Normally, in talk of collections, it is taken for granted that the members have their identity independently of the collection to which they belong, and this presupposition I will stick to, too. In this and the next section, I will discuss only collections whose members are spatiotemporal particulars, i.e., ordinary things, events, processes, states of affairs, tropes, and whatever.

From where does the identity of a first-order collection of spatiotemporal particulars stem? If one thinks of a collection such as the collection of *all* water molecules, it may seem simple. The identity stems from the nature of the members of the collection. But such an answer cannot possibly be the right one when it comes to a collection of, say, 100 000 water molecules. There are many such collections, and their members are both of the same nature and are equally many. Therefore, the numerical identity of the collections must be given in some other way than by the nature and number of their members. The same conclusion is even more obvious in the case of collections defined by enumeration. For instance, the members of the mentioned collection of a chair, a flower, and a stone have no nature in common. With the identity question now posed—and kept in mind in the background—a brief detour to language and logic will, I think, be illuminating.

In this century's philosophy, the term 'plural' has come to figure prominently in quite a new way. Here are three book titles (the asterisks are added): *Logic & Natural Language: On \*Plural Reference\* and Its Semantic and Logical Significance* (Ben-Yami 2004), *\*Plural Predication\** (McKay 2006), and *\*Plural Logic\** (Oliver & Smiley 2013). And the last two authors have rightly noted and remarked: "Simons was a pioneer of plural logic" (ibid., 98). In Simons' mentioned 1982-papers, there is a connection between his views on natural numbers and his stress on terms with plural reference ('we', 'they', 'these Xs', 'the Xs', 'a, b, and c', etc.);<sup>12</sup> and I will add some observations on this relationship.

Simons' view about numbers is that numbers are certain *non-distributive* formal properties of multitudes (collections). Therefore, let's look at a non-distributively predicative sentence such as 'They are five'; meaning there are five members in a certain collection. The sentence contains a plural subject term ('They') and a non-distributive predication of the referent ('being five'). Predications containing numerals, however, are by no means the only kind of non-distributive predications. Oliver and Smiley claim that there are at least nine different kinds of one-place collective predicates, of which the kind 'number' makes up only one; even if they place it in "pole position" (ibid., 115–7). I want, and need for my purposes, to point out that in an important respect numerical predications differ from many other kinds of non-distributive predication. I will use a quotation from McKay's book as my point of departure. He writes:

For example, any of these [collective and non-distributive] predicates might be true of some people without being true of any one of them:

- They are shipmates (classmates, fraternity brothers)
- They are meeting together
- They lifted a piano
- They are surrounding a building
- They come from many different countries
- They weigh over 500 pounds

Such predications are a routine part of ordinary language use, yet standard systems of first-order logic provide no place for such non-distributive predication. (McKay 2006, 1)

Out of context, plural subject terms (e.g. 'they') always look as if they refer to a collection. In the numerical predication 'They are five', the predication

---

12. The same stress on plural reference is also to be found in the paper (Yi 1998) that is mentioned in footnote 1.

does not change the collection-nature of the referent of ‘they’, but this is what happens in McKay’s six chosen predications. When the persons referred to by ‘they’ are said to be shipmates or meeting together (the first two examples), they are no longer seen as merely non-related members of a collection; they become related to each other and become, at least linguistically, parts of a state of affairs.<sup>13</sup> Analogously, when the persons referred to by ‘they’ lift a piano or surround a house (the next two examples), they become regarded as functional parts of a temporary functional unit, and, consequently, are no longer a mere collection. Being regarded as coming from different countries (fifth example) is a bit more complicated, but it is also a case of becoming regarded as parts of a complex state of affairs. Weighing over 500 pounds together means that a property of each member is added together, which, once again, means that after the predication the referents of ‘they’ can no longer be seen as only a collection; they do now also appear as constituents of a state of affairs.

Put briefly, in McKay’s sentences the plural subject term refers to a number of entities that are parts of either a state of affairs or a functional unit. But neither states of affairs nor functional units are collections; both kinds of entities are at least in appearance more than a collection. As said, out of context plural subject terms look as if they refer to a collection, but in context—and that is the only place where they have real referents—they need not necessarily do so.

It is part and parcel of non-distributive predications that the subject term in the sentence in question becomes treated as referring to a unity, and I am by no means denying this. I am stressing two other things. First, I am pointing out the fact that non-distributive logical predicates such as ‘being shipmates’, ‘meeting together’, ‘lifting together’, ‘surrounding’, ‘coming from many countries’, and ‘together weighing over 500 pounds’ present the persons referred to by the word ‘they’ as being parts of a unity that is presented as, so to speak, being a *new entity* constituted by the collection at hand. Second, I am pointing out that this is *not* the case in a non-distributive numerical predication such as ‘They are five’. Here, the referents of ‘they’ are just as in McKay’s examples treated as a unity, but nonetheless there is no new entity that is constituted by the collection. In numerical predications, the non-distributive predication unites (as in all such predications) the collection, but only in the sense that

---

13. I write ‘at least linguistically’, since Simons has argued that, ontologically, there are no mind-independent states of affairs (Simons 2009). I am of the contrary view (Johansson 2004, ch. 3).

the collection becomes no longer treated only as many, but treated also *as one*.

My view can be put like this: In a sentence such as ‘They are five’, (i) the word ‘they’ highlights the collection *as many*, (ii) the non-distributive numerical predication ‘are five’ treats the collective referent of ‘they’ *as one*, and (iii) the sentence as a whole is about the collection as one-and-many.

Plural predications involving numerals are non-distributive, and they presuppose that a collection apart from existing as many exists also as one. This fact is the linguistic side of the observation I made early in this section: collections can be counted, and must therefore in some sense be not only collections as many, but also collections as one. When two collections are equinumerous, each *as one* has *as many* the same number of members as the other. Collections should be regarded as one-and-many.

##### 5. *On the nature of collection boundaries*

I can now return to the question from where the identity of a finite first-order spatiotemporal collection stems. Since every finite spatiotemporal entity must in some sense have a boundary that demarcates it from its surrounding, the question is from where the boundaries of finite spatiotemporal collections stem; I am not, remember, concerned with the boundaries of their members.

There are several subtle philosophical questions involved in the notion of a boundary (Varzi 2008), but in order to discuss the boundaries of collections all of them need not to be dealt with.<sup>14</sup> From Varzi’s article I pick the distinctions (soon to be explained) between (i) owned versus unowned boundaries and (ii) natural versus artificial boundaries.<sup>15</sup> Then I will add a distinction between (iii) manifest and virtual boundaries.<sup>16</sup> The important conclusion will be this: *all finite spatiotemporal collections have unowned and artificial boundaries that only need to be virtual*.

---

14. Simons has written a paper about boundaries, too (Simons 1991), but it is not relevant for my discussion.

15. The latter is often discussed also under the labels *bona fide* versus *fiat* boundaries (Smith & Varzi 2000).

16. Normally, the contrary opposite of ‘manifest’ is ‘latent’, and the opposite of ‘virtual’ is either ‘real’ or ‘true’, but for various reasons I think the opposition ‘manifest’ vs. ‘virtual’ creates the most fruitful connotations in the context at hand.

For unreflective common sense, probably, a boundary is primarily an entity between two entities; for instance, the black line that delimits the white surface within this square  $\square$  from its white surrounding. Such boundaries are by definition *unowned* boundaries. But, philosophically, they of course give rise to the question what the boundaries between the first boundary (the line) and what is inside and outside of it are like. In relation to solid material things, reflections on perceptions easily lead to the view that the boundary of a thing must be the utmost part of it; or, as Euclid and Aristotle said, a boundary of a thing is the “extremity” of the thing (Varzi 2008, sect. 1). This view might well, I think, be called the view of reflective common sense. According to it, a true boundary is always an *owned* boundary; the boundary belongs to and is a kind of part (perhaps infinitesimal) of the entity it delimits. However, in mathematics curious entities have seen the light. For instance, the boundaries of an open interval are necessarily external to the interval; they are *unowned* without there being anything between the interval and the boundary. The open interval of points between the numbers  $a$  and  $b$  on the number line cannot but be defined by letting the external points  $a$  and  $b$  be the boundaries, i.e., the interval is defined as the set of points  $\{x \mid a < x < b\}$ .

Look at the following seven spots:  $\bullet \bullet \bullet \bullet \bullet \bullet \bullet$ . Among them, in the middle, there is a collection of *five* spots. The boundary of this collection cannot come from any of the five members alone, and to say that it comes from them collectively makes it a magic how these can make the exactly similar spots to the left and the right of the central five-spots collection to be outside of the collection. The boundary of the five-spots collection cannot be found within the collection, and must therefore be unowned. As far as I can see, the nature of the case makes this true of all finite collections. *All finite collections have unowned boundaries.* This is due to the fact that members of a collection have their individual identity independently of the collection to which they belong.

A *natural* boundary is a boundary grounded in some kind of pre-given discontinuity or gap between an entity and its surrounding. An *artificial* boundary is a correspondingly decision-grounded man-made discontinuity or gap. The perceptual system creates of itself an immense number of natural boundaries, which make it contain many discrete entities. A philosophically perplexing such boundary is that between two colored surfaces, e.g. the boundary between this black spot  $\bullet$  and its white surrounding. It is perplexing in the sense that it is hard to become clear about whether

the boundary is black (and owned by the spot), white (and owned by the surrounding), black-and-white (owned by both), or colorless (unowned by both). Surely, however, there must in some way or other be a boundary that makes the black spot distinct from the white surrounding. *Within* the homogeneously colored black spot, on the other hand, all the possible parts need an artificial boundary in order to come into existence. Quite generally, parts of homogeneous spatial wholes have to have their boundaries artificially created; but what to say about collections?

We can easily talk about collections such as ‘the pebbles in the bowl’, where the bowl functions as the unowned boundary of the collection. Since the bowl is man-made, it is an artificial boundary. The collection, however, would retain its identity even if the pebbles would be put in a natural pit. Since the boundary of the collection is unowned, it doesn’t matter what kind of material thing we decide to take as its unowned boundary. As noted, the boundary can be changed, but the collection remains the same. Therefore, even the natural pit mentioned must be regarded as an artificial boundary of the collection. It is a man-made choice that makes the collection to be confined within the pit, and which does not allow it to have any pebble outside of the pit as a member. As far as I can see, the nature of the case makes this true of all finite collections. *All finite collections have artificial boundaries.*

Artificial boundaries can be either manifest or virtual. With a *manifest* boundary I mean a boundary that is either completely grounded in something mind-independent or grounded in discontinuities in perception. Man-made material walls, as well as lines on a map, can be manifest boundaries; also, both this perceived square  $\square$  and this perceived spot  $\bullet$  have manifest boundaries. A good example of a *virtual* boundary is the boundary between two pieces of land before it has been marked off on either the ground or on maps, but nonetheless talked about and decided on.

In order to individuate a collection privately, it is enough to imagine for oneself the members as together having a boundary, and thereby constituting a collection, but in order to individuate a collection socially, one has to be able to talk about the delimitation in some way or other. It is only the latter kind of boundary that is of interest for this paper, since we are concerned with numbers within mathematics understood as an intersubjectively communicable discipline; and such boundaries need to be at least virtual boundaries. All private and virtual boundaries are artificial boundaries, but not vice versa.

When qualifying a noun, the term ‘virtual’ means being very nearly the kind of thing mentioned by the noun; implying that the differences are not especially important. For example, to have a virtual monopoly is to have a kind of monopoly, but not a true monopoly. In optics, ordinary mirror images are called virtual images. Such images can easily be seen, but they cannot be obtained on a screen; i.e., they can be seen but not touched, but it is the seeing that is important. In computer technology, virtual realities are simulated environments that to a person’s consciousness seem to be real or almost real. What I call a virtual boundary is a boundary that can neither be touched nor seen, but nonetheless has an intersubjective reality.

A virtual boundary is a social fact. When saying so, I connect to John Searle’s analyses of such facts (Searle 1995, 2010). He claims that the logical structure of institutional social facts that are grounded in material entities can be captured by this simple expression: *X counts as Y in C*.<sup>17</sup> In more detail it looks like this: *the material entity X counts as having the status function Y in the context or collective intentionality C*. Let me explain by means of an example.

Coin money fits the formula nicely. A ten euro coin is a piece of metal (X) that counts as, has the status of, and functions as ten euro (Y) in many European countries and in international business (C). In C, specific pieces of metal and paper are both described and perceived—with far-reaching consequences—as money. In a mind-independent world, the same pieces would have much more limited effects. From the point of view of nature, money is only a fiction or a simulation, but from the point of view of society, money is as real as anything can be. When a material thing counts as money in a community, it *is* money in the community; it is what it is regarded as.

In my opinion, the same kind of analysis can be applied to collections. If the pebbles  $p_1, p_2, p_3, p_4,$  and  $p_5$  count as a collection in a certain community, then here they are a collection. If the virtual boundary that makes them a collection counts as a boundary, then it is a boundary. *Collections can be just as real as money; and, from the point of view of nature, just as unreal*. As coins have one foot in nature and another in the social realm, collections of material entities have one foot in nature and another in the social realm. The members of a collection are material entities with presum-

---

17. I have in the review (Johansson 2011a) and in a number of other papers discussed Searle’s philosophy.

ably material boundaries, but the boundary of the collection is artificial and has at least virtual existence.

This was my brief defense of the claim that *all finite collections have unowned and artificial boundaries that only need to be virtual*. Simons, however, treats his multitudes as if there is no problem with how they are individuated. From my perspective, thereby, he implicitly treats spatiotemporal multitudes/collections as if they have *owned and natural boundaries*. As I will show, this difference is of consequence. Let it only be added that I find it understandable that unowned virtual boundaries become neglected. Since they are unowned they do not belong to the entities they delimit, and when they are virtual they are from the point of view of nature fictions.

Simons says that “multitudes [collections] with the same members are identical” (Simons 2011, 5), and with this I do of course agree. But from this fact he draws the false conclusion that “Multitudes are purely extensional collections” (ibid.). On my analysis, they are not *purely* extensional; they have, and necessarily so, also at least a virtual boundary as a condition for their existence. They require, pace (Simons 2011, 17), more for their existence than the existence of their members.<sup>18</sup> Note, though, that no *specific* boundary can be an identity condition of a collection. Since the boundary is unowned, a collection can retain its identity even if the boundary is exchanged. I have already mentioned that one and the same collection of pebbles can have either a man-made bowl or a natural pit as its boundary. Now it is time to add that a collection can retain its identity even if a manifest boundary (e.g., the bowl or the pit) is exchanged for one that is only virtual.

Common sense does not bother about the ontological status of the boundaries of collections, and in everyday life people take it for granted that we can create collections of arbitrary kinds of entities. That is, it is taken for granted that the human psyche/brain contains some kind of faculty or mechanism of imagination whereby collections can be freely created. Of course, philosophers who accept the existence of such a faculty, e.g. Husserl, come to the same conclusion. Even Hume, who denied the existence of enduring egos, but nonetheless let a faculty of imagination play

---

18. I have argued that collections are not self-individuating entities, but I am of the same opinion in relation to states of affairs. An artificial boundary is needed in order to isolate one finite state of affairs from all the encompassing state of affairs. For more reflections on this view and its repercussion on views in cognitive linguistics, see (Smith 2000).

a large role in his philosophy, is of this opinion.<sup>19</sup> In *Treatise* (1.2.2.3) he distinguishes between true unities (simple impressions and simple ideas) and fictional unities. He writes:

'Tis evident, that existence in itself belongs only to unity, [...] For by the same rule these twenty men *may be consider'd as an unite*. The whole globe of earth, nay the whole universe *may be consider'd as an unite*. That term of unity is merely a fictitious denomination, which the mind may apply to any quantity of objects it collects together. (Hume 2000, 25)

Put in other words: any number and kinds of existing true unities that the mind collects together might be called a unity, but in many such cases the new unity must be considered fictitious. In those cases, put in my words, there are collections of real particulars, but the boundaries are artificial.

Epistemologically, my analysis of collections implies that a collection cannot be truly apprehended in any other way than as simultaneously being both a unity (a collection as one) and a plurality (a collection as many), i.e., as unity-and-plurality, as one-and-many. When the unity is in the foreground, the plurality is in the background, and vice versa. But both of them have to be there all the time.

Mostly, singleton sets are regarded as being distinct from their member, i.e.  $\{x\} \neq x$ . The view of collection boundaries that I have put forward affords a similar answer to the question whether one-membered collections should be regarded as being distinct from their single member. Simons says that he regards an individual “as a degenerate case of a manifold [multitude, collection]” (Simons 1982a, 186), and that he will “stretch the meaning” of the term ‘multitude’ to allow a single individual to be a multitude [collection] (Simons 2011, 6). On my analysis, there is a clear difference between the member of a one-membered collection and the collection: *the member and the collection have numerically different boundaries*. Moreover, these boundaries are also of different *kinds*.

Look at the perceptual unity of the black spot •. As a unity it has a natural boundary (be it owned or unowned), but as a one-membered collection it has another kind of boundary, namely an artificial *virtual* boundary. Even a collection with only one member has a boundary that differs from the boundary of its single member. That is, I agree with Simons that there are one-membered multitudes [collections], but contra Simons I think there

---

19. In (Johansson 2012) I have shown the central role that the faculty of imagination plays in Hume's philosophy.

is no need here to talk about “degenerate cases” or “stretched meanings,” even if in everyday language no one talks about one-membered collections.

My general claim is that collections, because of their specific kind of boundary, can and should be described as being one-and-many. In the next section, the expression ‘one-and-many’ will be specified into expressions such as ‘one-and-7’ and ‘one-and-5’, which refer to collections with 7 and 5 members, respectively. In the same way, a single-membered collection can be referred to by the expression ‘one-and-1’.

## 6. *Higher-order collections and arithmetic operations*

Assume we have four collections, one with 2 members (one-and-2), one with 3 (one-and-3), one with 4 (one-and-4), and one with 5 (one-and-5). Connecting to a symbolism used in (Simons 2011), I will symbolize them  $[e,e]$ ,  $[e,e,e]$ ,  $[e,e,e,e]$ , and  $[e,e,e,e,e]$ , respectively. Here, the symbol ‘e’ is shorthand for the term ‘entity’, and the square brackets symbolize—for me, but not for Peter Simons—the boundary of the collection in question. Out of these four collections we can easily create two *second-order* collections; one consisting of the first two first-order collections, and another of the last two collections. Thereby we obtain:  $[[e,e] , [e,e,e]]$  and  $[[e,e,e,e] , [e,e,e,e,e]]$ . We can equally easily create a third-order collection of the two second-order collections:  $[[[e,e] , [e,e,e]] , [[e,e,e,e] , [e,e,e,e,e]]]$ . In other words, as there can be collections, there can be collections of higher order. As Husserl says:

Consequently it is a fact that we have the capability of representing several totalities together as unified into *one* totality, without thereby their separate unifications being lost. We represent totalities whose elements are in turn totalities. In fact, even totalities of totalities of totalities are thinkable, etc. (Husserl 2003, 97)

From the views defended in the former section it follows, that independently of whether a collection is a first-order or an  $n$ -order collection, its constituting boundary is unowned and needs only to be virtual. The fact that the boundaries have such a character means that they can never turn the contained multitude/collection into a new kind of entity. That is, even higher-order collections of concrete entities are in a sense only concrete collections; something which ought to please Simons. He says, as earlier quoted: “multitudes, of whatever order, are concrete if their members are

concrete.” From my own perspective, it even looks as if I have filled in a lacuna in Simons’ multitude theory, namely supplied an answer to the question how higher-order multitudes are constituted.

The really important point, however, is not the fact that we can form collections of collections of collections as we can form sets of sets of sets. The point to stress is that we can *simultaneously* cognize more than one collection level in a higher-order collection, and see that they are included in each other. Many-leveled apprehensions are well known and regarded as important in other areas, but as far as I know never stressed in the philosophy of mathematics. For instance, in fractal art, part of the seeing consists in seeing that the same geometrical shape comes back again and again with ever decreasing size; the shapes are contained in one another in something like the way the dolls of a Russian matryoshka doll are contained within one another. In conversation, verbal irony is necessarily a two-leveled apprehension. When saying the opposite of what one really means, one’s listeners must understand that there are two layers of meaning, an explicit surface meaning plus an implicitly present antithetical meaning that is the true message.

Such a view of many-leveled apprehensions of higher-order collections opens up new possibilities about how to look upon the distinction between addition and multiplication; first some words about addition and then about multiplication.

According to the traditional view, multiplication of integers can always be regarded as a more or less complex procedure of repeated addition; and, surely,  $3 \times 4 = 3 + 3 + 3 + 3$ . However, in the empirical science of number cognition, the view that multiplication must always be understood as repeated addition has been questioned. Here are two quotations:

Their [the children’s] performance could not be explained by processes of repeated addition, [...] These findings provide evidence for an untrained, intuitive process of calculating multiplicative numerical relationships. (McCrink & Spelke 2010, 204)

The hypothesis of conceptual discontinuity between additive and multiplicative reasoning receives further support from the fact that the children in the correspondence group made significantly more progress in multiplicative than in additive problems. (Park & Nunes 2001, 771)

Now, of course, it might be claimed that these empirical findings about how the brain works cannot falsify the view that, *from a logical point of view*, the operation of multiplication must be understood as being based

on that of addition. With this claim I agree, but if it can be shown that also from a logical point of view multiplication *can* take place independently of addition, fundamental things have to be re-thought. And I think the existence of second-order collections makes a kind of multiplication without addition possible and logically understandable. I am by no means discussing what may correspond to multiplication processes in brains and computers.

On the view I am expounding, remember, there can be no natural numbers if there is no prior cognitive activity that creates boundaries that delineate collections as one-and-many. On this presupposition, the natural view of addition involves only first-order collections, and it looks as follows.

Take Kant's old example of an arithmetic incontestable truth:  $7 + 5 = 12$ . Below is a list of four equality sentences related to that truth; the first belongs to pure arithmetic and the others to the highest abstract level of applied mathematics:

- a)  $7 + 5 = 12$
- b)  $7 e + 5 e = 12 e$ , or: one-of-7 + one-of-5 = one-of-12
- c)  $[e, e, e, e, e, e, e] + [e, e, e, e, e] = [e, e, e, e, e, e, e, e, e, e, e]$
- d) a collection of 7 entities + a (distinct) collection of 5 entities = a collection of 12 entities

Let me explain the list bottom up. In all cases the addition symbol '+' represents the operation of combining the two (explicitly or implicitly mentioned) collections on the left hand side to a new collection that is placed on the right hand side. From my perspective, addition is a sort of *boundary-reconstruction*.

I regard sentence (c) as merely a symbolization of sentence (d), and sentence (b) as merely a symbolization of sentence (c). A Platonist mathematician would regard (b) as an application of a truth about mathematical numbers, but according to the property view of numbers, (a) should be understood as being an abstraction from the truth stated by (b).

Only collections of things of the same kind, i.e., collections whose members are described by the same count noun, can be added. For example, (a collection of) 7 apples + (a collection of) 5 apples = (a collection of) 12 apples, but (a collection of) 7 apples + (a collection of) 5 bananas have no meaningful common sum; it is only a collection of seven apples and five bananas. However, the sum 7 fruits + 5 fruits = 12 fruits is as meaningful

as that concerned with apples, and when using the most abstract count noun possible, 'entity', we can add collections of whatever kind of phenomena we want. 7 entities (whatever their concrete nature is) + 5 entities (whatever their concrete nature is) = 12 entities.

Of course, what I have said about the addition  $7 + 5 = 12$  (or:  $7 e + 5 e = 12 e$ ) can be repeated in relation to the addition  $3 + 3 + 3 + 3 = 12$  (or:  $3 e + 3 e + 3 e + 3 e = 12 e$ ). Out of the four first-order collections on the left-hand side, which have 3 members each, a single first-order collection of 12 members is created. It is now a matter of *repeated boundary-reconstructions*. In the square bracket collection symbolism introduced, it can be written:

$$[e,e,e] + [e,e,e] + [e,e,e] + [e,e,e] = [e,e,e,e,e,e,e,e,e,e,e,e].$$

Now comes my thesis, namely that the multiplication  $3 \times 4 = 12$  need not necessarily be regarded as a truncated description of the addition procedure  $3 + 3 + 3 + 3 = 12$ , which contains only first-order collections. *It can be, no doubt, but it need not be.* If we allow the possibility of two-leveled apprehensions of second-order collections, we can make an immediate one-stroke cancellation of all the boundaries of the first-order collections, but nonetheless still have a collection. The second-order collection simply collapses into a first-order collection. When we apprehend a second-order collection that includes the apprehensions of four first-order collections, we may immediately see:

$$[[e,e,e] , [e,e,e] , [e,e,e] , [e,e,e]] = [e,e,e,e,e,e,e,e,e,e,e,e].$$

That is, a second-order collection with 4 first-order collections, each of which has 3 members, can immediately be seen to have as many members as a first-order collection with 12 members. It seems to me as if this equality can be regarded as logically representing multiplication without addition. It is as possible here to go directly (by means of boundary cancelling) from the single *second-order* collection on the left-hand side to the single *first-order* collection on the right-hand side, as it is in the addition case to go (by means of repeated boundary reconstructions) from four first-order collections to a single first-order collection. When one apprehends the second-order collection, one apprehends *both* the second-order collection *and* the contained first-order collections. To apprehend a second-order collection is to have a two-leveled apprehension.

On this analysis, both multiplication as repeated adding and multiplication without adding contain boundary-reconstructions, but the reconstruction operations are of different kinds.

The positing of apprehensions of levels within second-order collections has also another interesting consequence. It does in a simple way explain the following fact, discussed in mathematical pedagogic (Yoshida 2009). In pure arithmetic,  $3 \times 4 = 4 \times 3$ , and there is no need to bother about what number is the multiplicand (first factor) and what is the multiplier (second factor). In applied arithmetic, however, things are different. If a child is told to calculate by multiplication how many apples there are in 4 baskets with 3 apples each, it should use the multiplication  $3 \times 4$ , three (apples) times four (baskets), but if there are 3 baskets with 4 apples each, it should use  $4 \times 3$ , four (apples) times three (baskets).

In the square bracket collection symbolism I have introduced, the difference mentioned is made visible as a difference between two different kinds of second-order collections. That is: the purely arithmetic expression ' $3 \times 4$ ' is abstracted from the applied arithmetic expression  $[[e,e,e], [e,e,e], [e,e,e], [e,e,e]]$ , but the purely arithmetic ' $4 \times 3$ ' is abstracted from the applied arithmetic expression  $[[e,e,e,e], [e,e,e,e], [e,e,e,e]]$ . These second-order expressions are *as such* distinct, below symbolized as  ${}_2 \neq_2$  (the subscripts indicate that only second-order entities are related), which means that we can state that

$$[[e,e,e], [e,e,e], [e,e,e], [e,e,e]] \neq_2 [[e,e,e,e], [e,e,e,e], [e,e,e,e]],$$

even though these two distinct second-order collections are grounded in one and the same first-order collection,  $[e,e,e,e,e,e,e,e,e,e,e,e]$ .

## 7. *Spatiotemporal collections and fictional collections*

One problem for the view that natural numbers are properties of collections of concrete entities is how to make sure that one can talk about infinitely many natural numbers. Peter Simons' present solution is that this becomes guaranteed by the existence of higher-order multitudes/collections (Simons 2011, 16f.). After  $n$ -order collections there are  $(n+1)$ -order collections, and so on. However, since I already think a faculty of imagination is necessary in order to create first-order collections, not to speak of an infinite number of higher-order collections, I think the simplest way

out is to posit *fictional collections* as bearers of the natural numbers of pure arithmetic. The members of such first-order collections can be regarded as being as eternal and unchangeable as Platonic entities are regarded by the Platonists. Some words about this possibility.

When numbers are talked or written about, it is normally taken for granted that different numerals such as the Arabic '3' and the Roman 'III' refer to the same number, and that number concepts belonging to different number systems that correspond to each other, such as 3 of the decimal system and 101 of the binary system, refer to the same number. It is taken for granted that whatever numbers are, they are not to be identified with the symbols and concepts used in order to refer to them.

The lone Arabic number concepts of the decimal system, 1, 2, 3, etc., cannot be used to refer to anything in the spatiotemporal world unless they are connected to a count noun or a word that functions as such. There are no pure numbers in the primordial perceptual field or in the mind-independent spatiotemporal world. This is shown by the fact that if one is asked to point at 5, one has to ask '5 of what?' If, then, the asking person says 'take flowers', one can point at a collection of 5 flowers; if the answer is 'take pebbles', one should point at a collection of 5 pebbles. However, whatever kind of concrete collection one points at, it will not be an eternal and unchangeable collection. This is to me the central reason why pure arithmetic with its seemingly eternal truths such as '7 is larger than 5' request either Platonist or fictional collection referents. Since I am a naturalist, I will only consider fictions.

Everyday life contains a clear distinction between real people of flesh and blood and purely fictional people presented in novels, movies, and TV-series. Only people of the first kind have a biological body that can become truly sick, truly suffer pain, and literally die. When people are held apart in this way, the fact-fiction distinction is unproblematic. However, in story-telling it can easily be made very fuzzy. In the literary genre of fictionalized autobiographies, so-called autofiction, it is impossible for all except the author to tell what is intended as describing real facts and as describing fictional facts about the main figures involved. If the reader cares, she has to make hypotheses about the autobiography the way natural scientists make hypotheses about nature. To propose a hypothesis is neither to assert that a certain presumed fact is real nor to assert that it is a fiction; putting forward a hypothesis is a specific kind of language act.

Our natural language allows us to distinguish between factual assertions, fictional assertions, and factual hypotheses. We are even capable of a fourth

kind of “aboutness” language act, namely the mere entertaining of facts. That is, we can talk about something without treating it as being real, fictional, or the content of a hypothesis. The semantic-philosophical question to ask is not whether these distinctions are illusory or not; the semantic-philosophical question to ask is how they are possible. This question, however, I will not touch upon here. In my opinion, the philosophy of numbers can rest content with a mere acceptance of the fact that quite often we can identify, re-identify, and communicate about certain fictions. Therefore, it is possible to ascribe fictions a certain mode of existence, even though it is a mode that make them at bottom partly mind-dependent (Johansson 2010). This kind of existence-fictionalism must not be conflated with what today in the philosophy of mathematics goes under the name of fictionalism; the latter regards mathematical entities as straightforwardly non-existing and, therefore, mathematical statements as being false (Balaguer 2008).<sup>20</sup>

From a linguistic point of view, the dictionary meanings of words are the same independently of whether the words appear in factual, fictional, hypothetical, or mere-entertaining discourses. Similarly, sentences that represent language acts such as orders, promises, exclamations, and declarations have a lexical meaning that is independent of kinds of discourse. These facts explain why no re-reading is necessary if one has read a story as being factually true, but later is told that it is a complete fiction. Also, conversely, we can read something as if it is a fictional novel, but later be told that it is a true story. In both cases, we can immediately without any re-reading switch and see the story in the new light. Let me call this capability the *fact-fiction switch ability* (Johansson 2013). Turning factual assertions into fictional ones is always possible, but, of course, turning fictional assertions into non-absurd factual ones cannot always be done.

Likewise, as far as I can see, numerical expressions have the same lexical meaning independently of what kind of discourse they appear in. The sentence ‘There are 5 pebbles’ has the same lexical meaning in a fictional novel as in a factual report, and the purely arithmetic sentence ‘ $7 > 5$ ’ has

---

20. This is not an exhaustive classification of what falls under the term ‘fictionalism’. For instance, in (Gullberg 2011) a “fictionalism with truth” (ibid., 151) is defended. It differs from my position in two main respects. First, she does not have a property view of numbers; the mathematical truths she argues exist are truths about structures. Second, she is relying on a modification of Carnap’s distinction between questions that are external and internal to given conceptual frameworks, whereas I think it suffices to distinguish between spatiotemporal and virtual-fictional entities. The obvious question to ask Carnap and Gullberg is what ontological status they ascribe the seemingly absolute conceptual framework in which the external-internal distinction itself is made and presented.

the same lexical meaning when it is intended to be about fictions as when it is intended to be about Platonic objects. Using the fact-fiction switch ability, we can turn all old talk about Platonic mathematical objects into talk about fictions; the problem is only how to retain in the change the view that there are arithmetic truths. In other words, the problem is how to deny the second horn in this naturalist-mathematical dilemma: *If the referents of pure number concepts are eternal and unchangeable Platonic objects, we have to reject naturalism; but if they are fictional objects, we have to reject the view that arithmetic contains truths.*

Of course, if we allow fictional stories to contain logical contradictions, then it is impossible to defend the view that the statement ' $7 > 5$ ' is true, since the contradictory opposite ' $\neg(7 > 5)$ ' can be true, too. Therefore, the view to be discussed is whether non-contradictory arithmetic can contain truths if it has fictional referents, i.e., if the natural numbers are regarded as fictional entities, or, more precisely, as properties of fictional collections. My basic view is this. I think there are irreducible relations and operations, and if such there are, then assertions about relations and operations between fictional entities take on a special character.

If there are relational necessities stronger than that of physical necessity, then these make the corresponding relational assertions true independently of whether the relata are real or fictional. Look at these two triangles,  $\blacktriangledown$   $\blacktriangledown$ , and call them  $L_1$  and  $S_1$ , respectively. Obviously,  $L_1$  is larger in perceived area size than  $S_1$ . It means that if two other perceived triangles exist that are exactly like these, then by some kind of strong necessity  $L_2$  is larger than  $S_2$ . And the same must hold even if  $L$  and  $S$  are merely fictional triangles in a non-contradictory fictional story. If the relational sentence ' $L$  is larger than  $S$ ' is true by strong necessity, then it doesn't matter whether the referents are spatiotemporal or fictional; or Platonic for that matter.

Likewise, if numbers are properties of collections, then relational sentences such as ' $7 > 5$ ' and ' $7 + 5 = 12$ ' are true independently of whether the collections that have the properties of being 7 and 5, respectively, are real or fictional. In spatiotemporally applied arithmetic, the natural numbers are properties of real collections of spatiotemporally finite and often changeable entities, but in fictionally applied arithmetic the natural numbers can be properties of collections of fictional eternal and unchangeable entities. This is my view.<sup>21</sup>

---

21. Let me add one thing. Since I have the general view that properties exist within a determinable-determinate distinction (Johansson 2002), I regard each specific natural number as a determinate of a determinable that might be called '(natural) number' or 'multiplicity'. As

## 8. *Natural numbers and real numbers—equinumerosity and equiproportionality*

Peter Simons has claimed that even if the discovery of Russell's paradox made Frege's analysis of the natural numbers impossible, one should note that if Frege's logicism is extended to the real numbers, then other seemingly insuperable problems pop up (Simons 1987). Therefore, one may legitimately ask Simons—and those who, like me, tries to develop his property view of the natural numbers—how to look at the real numbers. I will conclude my paper with some remarks on this problem.

The property view of the natural numbers is tied to the notion of collections (or some synonym); and a collection is a collection of discrete entities. This means that the existence of the irrational numbers constitutes a problem for this approach just as much as it once did for the Pythagoreans. Of course, not even the rational numbers can when defined as fractions be monadic properties; they must be relations. However, since they can be regarded as relations between the number properties of two collections, they can at least in a simple way be grounded in monadic properties, but irrational numbers cannot. They require for such a grounding very complex mathematical operations; be it by Dedekind cuts, Cauchy sequences, or infinite series of nested fractions.

In order to illustrate the geometrical meaning of irrational numbers, I will use the two classic examples:  $\sqrt{2}$  and  $\pi$ . The fact that the irrational numbers cannot in a simple ratio way be defined by rational numbers means that there is no extended entity (measurement unit), however small, that can make up a concatenated collection on both the hypotenuse ( $= \sqrt{2}$ ) and the other sides ( $= 1$ ) of an isosceles right-angled triangle; and there is no extended entity, however small, that can make up a concatenated collection on both the circumference of a circle ( $= \pi d$ ) and its diameter ( $= d$ ). If the measurement unit searched for covers one of the lines exactly, it cannot possibly cover the other; and vice versa. How, then, in what way

---

red is a determinate of the determinable color, five is a determinate of the determinable number. If Husserl had made a distinction between determinable-determinate relations and genus-species relations, I think he would have put forward the same view. My reason is that the early Husserl can at least once be interpreted as claiming that the specific natural numbers are species of the genus cardinal number (Husserl 2003 [1891], 86), and that later in *Ideas* (Husserl 1982 [1913]) he becomes quite explicit about it: "cardinal number as such ["Anzahl überhaupt"] is a highest genus. Two, three, etc. are its *infimae species*." Unfortunately, the English book here contains a misleading mistranslation. "Anzahl überhaupt" is translated into "any cardinal number whatever" (Husserl 1982, 25).

should the property view of the natural numbers best be combined with a reasonable analysis of the real numbers? Simons says:

The relationship between ratios of quantities of like kind and the real numbers with which they can be uniquely *correlated* [italics added] is many to one: the numbers are, *qua* objects of pure mathematics, different from all their many applications. Euclid in fact worked *solely* with ratios: it took until the 19th century with the work of Dedekind, Cantor and others for the pure mathematical theory of real numbers to be established. (Simons 2013, 532)

If such a correlation view were to be applied to the natural numbers, it would claim that the relationship between collections and the natural numbers is merely one of correlation, not that the numbers are properties of collections. This runs completely counter to the property view approach. Therefore, I think that this approach should not too quickly be combined with a correlation approach to the real numbers. If the natural numbers are properties of collections, then this fact ought in some way or other to spill over on what the real numbers are; or so I think. If the ontological nature of the natural numbers cannot be understood merely by means of collections and numbers-as-objects, then something like a wide property view ought to be true of the real numbers, too.

Simons' property view of the natural numbers says that such numbers are non-distributive formal properties of multitudes/collections. And I have defended it by adding to Simons' position the notion of imposed collection boundaries and by accepting fictional collections. My proposal for the real numbers I will call *the proportionality view of the real numbers*. It claims (remember that 'formal' means being able to be meaningfully ascribed to different kinds of ontological entities):

*Real numbers are non-distributive formal proportion-relations between properties.*

Using my two examples, and leaving for the moment the term 'proportion-relation' unexplained, it amounts to the following.  $\sqrt{2}$  is a proportion-relation between the length of the hypotenuse and the length of the other sides of an isosceles right-angled triangle, be this triangle physical, mental, abstract, or fictional. Similarly,  $\pi$  is a proportion-relation between the length of the circumference and the diameter of a circle, be this circle physical, mental, abstract, or fictional. However, since both  $\sqrt{2}$  and  $\pi$  can be proportion-relations between many other pairs of properties, even

here Husserl's abstraction into "somethings" is meant to have come into play; this fact is symbolized by the term 'formal' in the expression 'formal proportion-relation'.

As it happens, my views have come to mind when thinking about two of Simons' papers and a very fine description of how physics became arithmetized; see (Simons 2013), (Simons forthcoming), and (Roche 1998, esp. ch. 3). They all take Euclid's treatment of ratios and proportions very seriously.

Even if Simons has never explicitly said so, it is easy to take him to mean that when we say things such as 'the wall is 107 m long', we regard the length of the wall *as if* it consists of a concatenated collection of 107 meter-length units.<sup>22</sup> The determinate length of the wall is in itself only a non-numerical property of the wall, and not a collection, but we can simulate collections in places where there are none. If we allow such *as-if-collections*, then many determinate magnitudes of determinable properties such as length, time duration, mass, and electric charge can be regarded as proportion-relations between two collections. In the case of length, one of the collections is then the collection consisting of the measurement unit for length, and the other is a simulated concatenated collection of such units that covers the magnitude measured.

We are working on the assumption that the natural numbers are properties of collections. Assume now that we have a collection (or as-if-collection) A that has the *property* of having 7 entities and another B that has the *property* of having 8. There are then *relations* between A and B, too. For instance, the sentences 'A has less members than B' and 'B has more members than A' are true. In my very considered opinion, these two sentences are about the same relation; the converse relation predicates 'have less members than' and 'have more members than' refer to the same relation. And the same thing can be said about a number of converse relation predicate pairs: 'smaller-larger', 'colder-warmer', etc. (Johansson 2014b). The view extends to purely arithmetic relations. The once famous mathematician and philosopher of physics Hermann Weyl writes in 1927:

Two propositions such as '5 follows upon 4' and '4 precedes 5' are expressions of one and the same relation between 4 and 5. It is unwarranted to speak here of two relations inverse to each other. (Weyl 1963, 4)

---

22. When it is a matter of other properties than that of length, then the term 'concatenation' has to be exchanged for others, but I will disregard all the measurement complications that are hidden here.

However, such relations, all of which might be called *more-and-less-relations*, are not what I will focus on; I have mentioned them only in order to show the plausibility of the view that converse relation predicates can describe the same relation. The kind of relations I want to zoom in on, are those that I have already talked about as being *proportion-relations*. Of course, where there is a proportion-relation there is a more-and-less-relation; but not necessarily vice versa.

Take a case where the collection A has to the collection B a proportion-relation of 7 to 8; or, using the converse relation predicate, that B has to A a proportion-relation of 8 to 7. I regard, as said, sentence pairs such as ‘A has to B the proportion-relation of 7 to 8’ and ‘B has to A the proportion-relation of 8 to 7’ to describe the same relation. Another way to put it is to say that ‘A and B has the ratio 7/8’, that ‘B and A has the ratio 8/7’, and that in these two sentences the two ratio numerals refer to the same proportion-relation.

The same proportion-relation can exist also as a proportion between collections with quite other numbers as *relata*. The ratio 7/8 is the same as the ratio 49/56, and the proportion-relation between two collections that have the number properties 49 and 56 is the same as the proportion-relation between two collections that have the number properties 7 and 8. If collection A has 7 entities, collection B 8 entities, collection C 49 entities, and collection D 56 entities, then (in Euclid’s very terms) A is to B as C is to D. I will call this (second-order) relation between the first-order proportion-relations ‘*equiproportionality*’.<sup>23</sup>

On the real number line, the ratios 7/8 and 49/56 are placed in exactly the same point. That is, the two ratios correspond to the same rational number.

Just as on the property view of the natural numbers an observed *equi-numerosity* is only a criterion for the state of affairs that two collections have the same number of members,<sup>24</sup> on the proportionality view of the

---

23. What I call ‘proportion-relation’, Euclid calls ‘ratio’; and what I call ‘*equiproportionality*’, he calls simply ‘*proportionality*’. As I see it, the term ‘ratio’ has today too strong connotations of necessarily being something arithmetic to fit my purposes; in Euclid’s time the term simply meant a kind of relation. Here is a sentence from Roche: “Some early Greek writers on mathematics, however, introduced the concept of the ‘number by which a ratio is called’” (Roche 1998, 46).

24. In his explicit critique of Frege (Husserl 2003, ch. VII), Husserl writes: “The definition of equivalence is, as we have shown, nothing more than a mere criterion for the existence of equality of number in two groups, whereas here [in Frege and some others] it is taken to be a nominal definition of it” (Husserl 2003, 121). Frege had very good reasons to write a long critical review of Husserl’s *Philosophy of Arithmetic*. He is explicitly under heavy attack in the book.

real numbers an observed *equiproportionality* is only a criterion for the state of affairs that two pairs of collections have the same proportion-relation.

So far, I have only made it clear that the rational numbers can be regarded as proportion-relations between properties, but the starting point of the investigation was the existence of the irrational numbers. So, what about them? Today, it may seem as if the mathematical constructions that place the irrational numbers on specific places on the real number line, and define them in terms of operations on the rational numbers, are so complicated, that it must be unreasonable to try to extend my view of the rational numbers to that of the irrational numbers. In my opinion, however, this view turns things upside down.

The remark made takes a purely arithmetic realm for granted, more precisely the real number line, and looks then for isomorphic correlations to other realms, but the actual discovery of the irrational numbers went the opposite way. It went from non-arithmetized geometry to arithmetic. It was discovered that the proportion-relations between the length of the hypotenuse (a property) and the length of the sides in an isosceles right-angled triangle (another property), and between the length of the circumference (a property) and the length of the diameter of a circle (another property), cannot be represented by any rational numbers and the then known arithmetical operations. Therefore, in whatever way it is nowadays in pure mathematics possible to define  $\sqrt{2}$  and  $\pi$  by means of operations on rational numbers, such constructions cannot show that the irrational numbers do not, like the rational numbers, refer to formal proportion-relations between properties.

One qualification, though, needs to be added; it was triggered by (Michell 1993). He works on a relational account of all numbers, even the natural numbers, and he has noted that proportion-relations are not in themselves necessarily uniquely determined. For instance (my example), the proportion-relation between the circumference and the diameter of a circle is  $\pi$  only in Euclidean geometry; in non-Euclidean geometries it takes on other values. To me, this does not mean that the real numbers cannot be regarded as being proportion-relations between properties; it means that proportion-relations might be unique only in relation to a pre-given theoretical system.

Obviously, it is as meaningful to talk about a proportion-relation between a collection with 2 members and a collection with 1 member, as it is to talk about a proportion-relation between a collection with 7 members and one with 8. But since in pure arithmetic  $2/1 = 2$ , the numeral '2' and

all other whole-number-numerals become in the approach I am pursuing ambiguous. When ‘2’ is used as referring to a natural number only among other natural numbers, it is used to refer to a monadic *property* of a collection, but when it is used as referring to a number on the real number line, it cannot refer to a property; it must refer to a *proportion-relation*. The belief that numbers primarily are defined by their place on the number line hides this ambiguity of the whole-number-numerals. When it is not noted, two kinds of mistaken ambiguity-cancelling can occur.

One mistake is to think that since the natural numbers are monadic properties, the same must be true of the integer rational numbers, too. The other is to think, conversely, that since all the rational numbers refer to proportion-relations, the same must be true of the natural numbers, too.

The first mistake occurs in lay people who unreflectively talk as if a thing that is, say, 1.75 m long, has this numerical value as an intrinsic property. They do not to themselves make it explicit that to be 1.75 m long is to be 1.75 *times* as long as a standard meter. Normally, this is an innocent mistake that does not affect what they are concerned with. The second mistake occurs in scientists and philosophers who always take the whole number line as their very starting point when thinking about mathematical physics. In the case I will mention, it is not an innocent mistake without consequences.

The International System of Units, the SI-system, is a very widely used system that fixes terminology and standard units for a large variety of measurements in physics and chemistry. Its history dates back to 1875 and the first international meter convention. In this system, the so-called base quantities are treated as if all number value ascriptions in physics and chemistry fit the structure of the real number line. Hereby, the distinction between discrete entities and continuous magnitudes is lost out of sight, which creates unnecessary complications in physical chemistry, where the chemists often use only formulas that contain natural numbers and symbols for count nouns naming kinds of atoms and molecules. In other words, measurement units and proportion-relations are introduced where natural numbers as properties of collections are enough. This criticism is put forward in (De Bièvre 2007, 2011, Price & De Bièvre 2009, Johansson 2011b, Cooper & Humphry 2012). I think that *one* explanatory factor behind the SI-system’s unhappy treatment of discrete entities is the ambiguity of the whole-number-numerals that I have tried to expose.<sup>25</sup>

---

25. Two other factors are, I am quite confident, (i) a misplaced longing to have one and the same mathematical structure for all physical-chemical quantities, and (ii) the fact that when scientists have to do with very large ensembles of discrete entities, it becomes mathematically

Before ending the paper, I would like to mention two similarities between the property view of the natural numbers, which brings in the notion of collection, and my proportionality view of the real numbers, which brings in the notion of proportion-relations. As said, the aim of this last section is only to sketch an analysis of the real numbers that has affinity with the property view of the natural numbers.

First, just as a pure natural number is a property abstracted from an entity collection, a pure real number is a proportion-relation abstracted from property pairs. For instance, as 5 is an abstraction from 5 entities, the real number  $7/8$  is an abstraction from a proportion-relation of 7 to 8.

Second, the property view of numbers is intimately connected with taking the existence of non-distributive plural predication logically seriously. Therefore, it should be noted that also sentences describing proportion-relations contain non-distributive plural predications. For example, in the sentence 'A and B have a proportion-relation of 7 to 8' the term 'A and B' is a plural subject, and the logical predicate 'having a proportion-relation of 7 to 8' is a non-distributive predicate.

\*\*\*

It would be nice if in the future the property view approach to the ontology of numbers can catch the minds of more philosophers than it has so far. Having said this, I end by thanking Peter Simons for having made at least me aware of this approach.

### *Acknowledgements*

I would like to thank Kevin Mulligan, Barry Smith, and Christer Svennerlind for comments on an earlier version of the paper.

---

practical for them to use the whole real number line *as if* there could be even an irrational number of discrete entities; (Johansson 2014c, sect. 7).

## REFERENCES

(All internet references retrieved December 2013)

- Armstrong, David M. 1978: *Universals & Scientific Realism I–II*. Cambridge: Cambridge University Press.
- 2010: *Sketch for a Systematic Metaphysics*. Oxford: Oxford University Press.
- Balaguer, Mark 2008: “Fictionalism in the Philosophy of Mathematics”. In: Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/fictionalism-mathematics/>. Stanford: The Metaphysics Research Lab.
- Bell, David 1990: *Husserl*. London: Routledge.
- Ben-Yami, Hanoch 2004: *Logic & Natural Language: On Plural Reference and Its Semantic and Logical Significance*. Aldershot: Ashgate.
- Boolos, George 1971: “The Iterative Conception of Set”. *The Journal of Philosophy* 68, 215–231.
- Bigelow, John 1988: *The Reality of Numbers*. Oxford: Clarendon Press.
- Bostock, David 2009: *Philosophy of Mathematics. An Introduction*. Oxford: Wiley-Blackwell.
- Campbell, Keith 1990: *Abstract Particulars*. Oxford: Basil Blackwell.
- Centrone, Stefania 2010: *Logic and Philosophy of Mathematics in the Early Husserl*. Dordrecht: Springer.
- Cooper, Gordon & Humphry, Stephen M. 2012: “The Ontological Distinction between Units and Entities”. *Synthese* 187, 393–401.
- De Bièvre, Paul 2007: “Numerosity versus Mass”. *Accreditation and Quality assurance* 12, 221–222.
- 2011: “Integer Numbers and their Ratios are Key Concepts in Describing the Interactions of Atoms and Molecules”. *Accreditation and Quality assurance* 16, 117–120.
- Euclid, 300 BC: *Elements*. <http://aleph0.clarku.edu/~djoyce/java/elements/bookI/bookI.html>.
- Forster, Thomas 2008: “The Iterative Conception of Set”. *The Review of Symbolic Logic* 1, 97–110.
- Frege, Gottlob 1972 [1894]: “Review of Dr. E. Husserl’s *Philosophy of Arithmetic*”. *Mind* 81, 321–337.
- Gullberg, Ebba 2011: *Objects and Objectivity. Alternatives to Mathematical Realism*. Umeå: Umeå Studies in Philosophy.
- Horsten, Leon 2012: “Philosophy of Mathematics”. In: Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/philosophy-mathematics/>. Stanford: The Metaphysics Research Lab.

- Hume, David 2000: "A Treatise of Human Nature". In: David F. Norton & Mary J. Norton (eds.), *David Hume: A Treatise of Human Nature*. Oxford: Oxford University Press, 1–417.
- Husserl, Edmund 2003 [1891]: *Philosophy of Arithmetic*. Dordrecht: Kluwer Academic Publishers.
- 1970: *Logical Investigations* vols. I [1913] and II [1921]. London: Routledge & Kegan Paul.
- 1982 [1913]: *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy (first book)*. Dordrecht: Kluwer Academic Publishers.
- Johansson, Ingvar 2002: "Determinables as Universals". *The Monist* 83, 101–121.
- 2004: *Ontological Investigations*. Frankfurt: Ontos.
- 2010: "Fictions and the Spatiotemporal World—in the Light of Ingarden". *Polish Journal of Philosophy* IV, 81–103.
- 2011a: "John Searle in the year 2010—Reviews". *Metaphysica* 12, 73–85.
- 2011b: "The Mole is Not an Ordinary Measurement Unit". *Accreditation and Quality assurance* 16, 467–470.
- 2012: "Hume's Ontology". *Metaphysica* 13, 87–105.
- 2013: "The Ideal as Real and as Purely Intentional—Ingarden-based Reflections". *Semiotica* 194, 21–37.
- 2014a: "Proof of the Existence of Universals—from a Fallibilist". In: Guido Bonino & Greg Jesson (eds.), *Defending Realism*. Berlin: De Gruyter, 51–61.
- 2014b: "On Converse Relations". In: Helge Malmgren & Torgny Nordin (eds.), *Philosophy and Botany*. Stockholm: Thales, 82–106.
- 2014c: "Constancy and Circularity in the SI". In: MetrologyBytes.net: <http://www.metrologybytes.net/opEds2014.php>
- Körner, Stephan 1960: *The Philosophy of Mathematics. An Introductory Essay*. London: Hutchinson.
- Lowe, E.J. 2006: *The Four-Category Ontology*. Oxford: Clarendon Press.
- McCrink, Koleen & Spelke, Elizabeth S. 2010: "Core Multiplication in Childhood". *Cognition* 116, 204–216.
- McKay, Thomas J. 2006: *Plural Predication*. Oxford: Clarendon Press.
- Michell, Joel 1993: "Numbers, Ratios, and Structural Relations". *Australasian Journal of Philosophy* 71, 325–332.
- Mulligan, Kevin 2013: "Formal Concepts". In: Kevin Mulligan, Katarzyna Kijania-Placek & Tomasz Placek (eds.), *The History and Philosophy of Polish Logic*. Basingstoke: Palgrave Macmillan, 205–223.
- Oliver, Alex & Smiley, Timothy 2013: *Plural Logic*. Oxford: Oxford University Press.
- Park, Jee-Hyun & Nunes, Terezinha 2001: "The Development of the Concept of Multiplication". *Cognitive Development* 16, 763–773.

- Price, Gary & De Bièvre, Paul 2009: "Simple Principles for Metrology in Chemistry: Identifying and Counting". *Accreditation and Quality assurance* 14, 295–305.
- Roche, John C. 1998: *The Mathematics of Measurement: A Critical History*. London: The Athlone Press.
- Russell, Bertrand 2006 [1903]: *Principles of Mathematics*. London: Routledge.
- Searle, John 1995: *The Construction of Social Reality*. New York: The Free Press.
- 2010: *Making the Social World*. Oxford: Oxford University Press.
- Simons, Peter 1982a: "Number and Manifolds". In: Barry Smith (ed.), *Parts and Moments*. Munich: Philosophia, 160–198.
- 1982b: "Plural Reference and Set Theory". *Ibid.*, 199–256.
- 1987: "Frege's Theory of Real Numbers". *History and Philosophy of Logic* 8, 25–44. Reprinted in Peter M. Simons, *Philosophy and Logic in Central Europe from Bolzano to Tarski*. Dordrecht: Kluwer, 1993, 117–142, and in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics*. Cambridge, Mass.: Harvard University Press, 1995, 358–385.
- 1991: "Faces, Boundaries, and Thin Layers". In: A.P. Martinich & Michael J. White (eds.), *Certainty and Surface in Epistemology and Philosophical Method. Essays in Honor of Avrum Stroll*. Lewiston: Edwin Mellen Press, 87–99.
- 1994: "Particulars in Particular Clothing: Three Trope Theories of Substance". *Philosophy and Phenomenological Research* 54, 553–575.
- 2005: "Against Set Theory". In: Johann C. Marek & Maria E. Reicher (eds.), *Erfahrung und Analyse/Experience and Analysis. Proceedings of the 2004 Wittgenstein Symposium*. Vienna: öbv&hpt, 143–152.
- 2007: "What Numbers Really Are". In: Randall A. Auxier & Lewis E. Hahn (eds.), *The Philosophy of Michael Dummett*. Chicago: Open Court, 229–247.
- 2009: "Why There Are No States of Affairs". In: Maria E. Reicher (ed.), *States of Affairs*. Frankfurt: Ontos, 111–128.
- 2010: "Foreword". In: Centrone 2010, v–viii.
- 2011: "On Multitudes". Lecture at Opole University: [http://www.gli.uni.opole.pl/publications/LT\\_2011\\_10\\_en.pdf](http://www.gli.uni.opole.pl/publications/LT_2011_10_en.pdf)
- 2013a: "Density, Angle, and other Dimensional Nonsense: How Not to Standardize Quantity". In: Christer Svennerlind, Jan Almäng, & Rögvaldur Ingthorsson (eds.), *Johanssonian Investigations*. Frankfurt: Ontos Verlag, 516–534.
- 2013b: "Arithmetic in Leśniewski's Ontology". In: Kevin Mulligan, Katarzyna Kijania-Placek, & Tomasz Placek (eds.). *The History and Philosophy of Polish Logic*. Basingstoke: Palgrave Macmillan, 227–241.
- forthcoming: "Euclid's Context Principle". *Hermathena—A Trinity College Dublin Review* (a Festschrift issue for Prof. John Cleary, 1949–2009).

- Smith, Barry 2000: "Truth and the Visual Field". In: Jean Petitot, Francisco J. Varela, Bernard Pachoud, & Jean-Michel Roy (eds.), *Naturalizing Phenomenology. Issues in Contemporary Phenomenology and Cognitive Science*. Stanford: Stanford University Press, 317–329.
- Smith, Barry, & Varzi, Achille 2000: "Fiat and Bona Fide Boundaries". *Philosophy and Phenomenological Research* 60, 401–420.
- Varzi, Achille 2008: "Boundary". In: Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/boundary/>. Stanford: The Metaphysics Research Lab.
- Weyl, Hermann 1963 [1927]: *Philosophy of Mathematics and Natural Science*. New York: Atheneum.
- Willard, Dallas 2003: "Translator's Introduction". In: Husserl 2003, xiii–lxiv.
- Yi, Byeong-Uk 1998: "Numbers and Relations". *Erkenntnis* 49, 93–113.
- Yoshida, Makoto 2009: "Is Multiplication Just Repeated Addition?" Lecture at 2009 NCTM Annual Conference: [http://www.globaledresources.com/resources/assets/042309\\_Multiplication\\_v2.pdf](http://www.globaledresources.com/resources/assets/042309_Multiplication_v2.pdf).

## IN DEFENSE OF SUBSTANCE

Kathrin KOSLICKI  
University of Alberta\*

### *Summary*

In his “Farewell to Substance: A Differentiated Leave-Taking”, Peter Simons reaches the provocative conclusion that the concept of substance, as it is employed by metaphysicians, has become obsolete, since in the end there may be nothing at all which answers to it. No harm is done, Simons allows, if we continue to retain an everyday notion of substance, as long as we are aware of the limitations of this practice: there is no reason in general to expect that what is salient from our specifically human point of view will retain a special place in light of our most considered scientific and metaphysical theories of the world. In this paper, I argue that, contrary to Simons’ pessimistic outlook, the concept of substance continues to retain its importance for metaphysics. Among the primary explanatory roles played by the concept of substance in metaphysics is its use in designating certain *kinds* of entities as occupying a privileged position relative to a particular ontology. But disputes over substancehood can also target the *criteria* themselves relative to which an ontologically privileged position is awarded to certain taxonomic categories. In these uses, we see the concept of substance employed in an absolute, a relational and a comparative sense, to designate items as substances simpliciter, as the substances *of* something else, or as *more* or *less deserving of substance status*.

---

\* I am very honored to be able to contribute to a volume which celebrates the work of Peter Simons. I have learned much of what I know about mereology from Simons’ monumental work, *Parts: A Study in Ontology* (Oxford: Clarendon Press, 1987), which I have come to regard as the “bible of mereology”. The painstaking efforts Simons took to familiarize the rest of us with foundational issues in mereology and their applications to central philosophical problems have paid off tremendously in placing questions concerning parts and wholes at the forefront of many current debates in metaphysics. Since it is an accepted practice in analytic philosophy to show one’s respect and admiration for a philosopher by subjecting his or her work to vigorous criticism, I follow this tradition here by giving Simons’ views concerning substance a run for their money.

## 1. *Introduction*

In his “Farewell to Substance: A Differentiated Leave-Taking”, Peter Simons argues for a provocative position concerning the usefulness and importance of the concept of substance for metaphysics:

Future metaphysics worthy of the name will need to be revisionary, and the concept substance will feature within it, at best, as a derivative construct. It is premature to say how such a future revisionary metaphysics will look, but it will need to both accommodate the advances of science as well as provide the platform for showing how we and our commonsense knowledge, including the knowledge of what have been thought of as individual substances, have a place within the same overall scheme. Substance will not be simply discredited, but its role as a fundamental metaphysical primitive is gone forever. Its formal moments, the notions of independence, of persistence, of unity and integrity, of discernibility, will need to be taken account of, but they will be analytical factors out of which the everyday notion is obtained, probably with some admixture of epistemological content in order to match the notion to its paradigm examples. (Simons 1998, 250)

Simons’ verdict is certainly surprising, given the centrality of the concept of substance across the history of Western philosophy. The notion of substance plays an important role throughout the ancient, medieval and modern period, in the works of philosophers as diverse as Aristotle, Aquinas, Descartes, Leibniz, Spinoza, Locke, Hume and Kant. Even in contemporary metaphysics, the concept of substance has experienced something of a revival, particularly in recent work by neo-Aristotelians and others writing on fundamentality, grounding, ontological dependence, essence, real definition and related notions.<sup>1</sup> It is therefore worth looking back to the considerations which prompted Simons to conclude that “there is no place for a fundamental concept of substance within [revisionary metaphysics], although aspects of the concept are likely to find their place therein” (*ibid.*, 235).

---

1. See for example Corkum 2008, 2013a, 2013b; Fine 1995; Gorman 2005, 2006a, 2006b; Gorman and Sanford 2004; Hoffman and Rosenkrantz 1994, 1997; Koslicki 2013a, 2015; Lowe 1998, 2006; Peramatzis 2008, 2011; Schaffer 2009; Toner 2008, 2010, 2011.

## 2. *Simons on substance*

Simons in his discussion speaks of at least two different ways in which the concept of substance can be employed, the first corresponding to what he calls the “humble” or “everyday substances” and the second to what he calls the “metaphysical substances”:

- (i) An ordinary everyday commonsense use of the concept of substance.
- (ii) A technical use of the concept of substance as it is employed by metaphysicians.<sup>2</sup>

Simons takes (i) to apply to certain kinds of macroscopic concrete particular objects which are accessible to us through our unaided senses and which figure prominently in our experience of the world. When Simons speaks of the everyday substances, he seems to have in mind roughly the primary substances of Aristotle’s *Categories*, e.g., organisms, other natural things which are not alive (e.g., atoms, mountains or planets), as well as artifacts:<sup>3</sup>

Material things, organisms, geographical features and heavenly bodies are our constant companions through life. We are born of them, marry them, make them, change them, destroy them, buy and sell them, explore them. We fill our waking and sleeping hours talking and thinking about them. (Ibid., 238)

---

2. In what follows, I will speak of different *uses* of the concept of substance, rather than of different *concepts* of substance. But I have no strong commitments concerning the individuation of concepts or their uses. My arguments could easily be reformulated in terms of a distinction between different substance concepts, whenever I refer to distinct uses of the concept of substance.

3. I am skeptical as to whether there really is such a thing as an ordinary everyday commonsense notion of substance of the kind Simons has in mind, in addition to the technical philosophical uses of the concept. Ordinary speakers of English certainly employ the term, “substance”, e.g., in such phrases as “illegal substances” or “banned substances”. But such non-philosophical uses of the term do not dovetail with the way in which the expression is employed by philosophers. One might think, of course, that ordinary speakers of English nevertheless are in possession of a *concept* of substance, which they employ in an everyday commonsense way, even if they do not use the expression, “substance”, to convey this concept. But it is not clear to me what sort of evidence would support this latter claim. When ordinary speakers of English employ more indiscriminate labels, such as “thing”, they use them to pick out a much wider range of entities than what philosophers have in mind when they assign the privileged position of substances to certain entities in their ontologies. For the purposes of this discussion, however, I will grant Simons’ assumption that the concept of substance has an ordinary everyday commonsense use, in addition to its use as a technical philosophical concept.

In Simons' view, (i), the concept of substance when used in the ordinary everyday way, is "perfectly harmless", as far as it goes, and does not need to be revised or discarded (ibid., 235). In contrast, Simons takes (ii) to be problematic and to have outlived its usefulness, due to a large extent to the following two assumptions which Simons associates with at least certain philosophical uses of the concept of substance:

- (a) The concept of substance must be accepted as a basic unanalyzable primitive.
- (b) The concept of substance applies to entities marked by seven characteristics: independence, ultimate subjecthood, individuality, persistence, referential salience, unity and integrity.

Given the complexity of the subject matter and the many treatments it has received over the years, Simons of course does not presume that the history of Western philosophy presents us with anything like a single unitary philosophical use of the concept of substance. Rather, he proceeds in his discussion by singling out several prominent strands, namely those stated in (b), which he takes to have been traditionally associated with at least certain philosophical uses of the concept of substance. The strands stated in (b), which Simons singles out for special attention, on the whole, fit reasonably well with Aristotle's conception of primary substance in the *Categories*, though even there the match is not perfect, since unity and integrity are not explicitly mentioned by Aristotle in the *Categories* as special marks differentiating the primary substances from everything else. (As we will see below, it is doubtful that Aristotle himself in the *Categories* or elsewhere would have accepted (a) as a constraint governing his philosophical uses of the concept of substance.) More generally, Simons' prominent strands certainly cannot be taken to be representative of how the concept of substance is used in other Aristotelian texts or by other philosophers. At best, then, (a) and (b) together only single out *a* metaphysical use of the concept of substance (or a family of such uses), and not *the* metaphysical use of the concept of substance (since there is no such thing). For the time being, I will follow Simons' usage and speak of *the* metaphysical use of the concept of substance as one which is associated with the constraints or desiderata stated in (a) and (b). I will, however, present reasons below which call into question the wisdom of this practice.

According to the prominent strands singled out by Simons, the substances are, first, *ontologically independent* beings, i.e., entities which are

in a certain sense self-sufficient. Simons construes the self-sufficiency or ontological independence in question as consisting in an entity's ability to exist while relying on nothing more than possibly the existence of its own proper parts. An individual human being, such as Socrates, in Simons' view, would count as ontologically independent in this sense, but the particular instance of paleness, for example, which inheres in Socrates at any particular time at which he exists, would not.<sup>4</sup> Secondly, the substances are *ultimate subjects of predication*. While it can be said of Socrates that he is pale, it can only be said of Socrates himself that he is Socrates (and here only in the sense of identity). Thirdly, the substances are *individuals*, rather than universals. Other things besides Socrates can be pale as well; but Socrates himself is not the kind of thing which can be shared between numerically distinct entities by being wholly present in different regions of spacetime. Fourthly, the substances are capable of *persisting through intrinsic change* over time. When Socrates goes from being pale at one time to not being pale at another, it is Socrates himself who undergoes intrinsic change with respect to certain of his characteristics. In contrast, if the belief that Socrates is pale goes from being true to being false, it does so merely as a result of the intrinsic change which takes place in Socrates. Fifth, the substances are preferred *objects of reference*: they occupy a privileged position with respect to our discourse, thoughts and actions. Natural languages, for example, reserve a proper name for Socrates, while paleness is standardly represented by means of a general term. Finally, the substances are marked by *unity* and *integrity*. To illustrate, Socrates, while he is alive, contrasts with the corpse he leaves behind in that the parts of the former are integrated into a living organism in such a way that they compose a unified whole, while the parts of the latter are only loosely assembled and slowly disintegrate into their surroundings.<sup>5</sup>

---

4. In this context, we have to exclude *causal* requirements, such as Socrates' need to breathe air filled with oxygen in order for him to survive. For reasons I have stated elsewhere (Koslicki 2013a, 2013b), Simons' construal of ontological independence as modal existential independence is not my own preferred notion. For one thing, modal existential independence does not yield the most plausible reading of Aristotle's conception of primary substance as developed in the *Categories*. In addition, modal existential independence turns out not to be the most fruitful notion from the point of view of formulating a plausible criterion of substancehood more generally. Since I have discussed these issues in detail in other work, it will not be necessary to dwell on them here. See also Corkum 2008, 2013a, 2013b and Peramatzis 2008, 2011, for further discussion.

5. Simons' work in mereology, especially Chapter 9 of Simons 1987, has contributed greatly to our understanding of integrity as a property of wholes.

Simons reaches the following overall verdict with respect to the everyday use of the concept of substance in (i). Insofar as we retain (i), we must acknowledge that the concept of substance, when used in the ordinary everyday commonsense way, does not satisfy all the desiderata cited above in (a) and (b). Simons therefore concludes that the everyday substances are not metaphysical substances; and the everyday use of the concept of substance is not what metaphysicians have in mind when they employ the notion in a technical role.

Simons' main reasons for thinking that the concept of substance, when used in the ordinary everyday commonsense way, cannot be construed as basic and therefore fails the first constraint in (a) are as follows. For one thing, according to Simons' preferred analysis (cf., Simons 1994), everyday substances turn out to be complex trope bundles of a certain kind, and are hence further analyzable into more basic constituents, viz., the tropes or particularized property instances which characterize these wholes. As a result, Simons takes the concept of substance in its everyday use to be a defined notion and therefore non-basic. Secondly, Simons notes that, in order to explain the characteristics and behavior of everyday substances, our best scientific theories (in particular, quantum mechanics) find themselves appealing to such entities as quarks, fermions and electrons which are not immediately accessible to us through our unaided senses. While these micro-physical entities play a central role in our comprehensive scientific understanding of the world, they do not figure directly into our commonsense representations.

In connection with the second constraint, Simons argues that the everyday substances do not satisfy all of the characteristics specified in (b) and hence are not to be identified with the metaphysical substances. While Simons is happy to ascribe *some* of the seven characteristics identified above to the everyday substances, he adduces both metaphysical and scientific considerations to the effect that these entities do not in general satisfy *all* of these characteristics simultaneously. In particular, the everyday substances turn out not to be the straightforwardly re-identifiable individuals metaphysicians might have hoped for.

On the positive side, Simons 1994 offers a careful treatment of independence, unity and integrity, according to which these properties can be successfully ascribed to everyday substances. In addition, given their prominent role in our experience of the world, Simons of course has no qualms granting that everyday substances are suitable to act as preferred objects of reference. Their status as ultimate subjects of predication is

furthermore unproblematic, since they are the bearers of properties (represented by the tropes in each bundle) but are not themselves properties exemplified by other trope bundles. The everyday substances, for Simons, thus satisfy at least five out of the seven characteristics he singles out for special consideration.

Metaphysical troubles do arise, however, when it comes to the ascription of the remaining two characteristics to the everyday substances: their alleged status as individuals, capable of persisting through intrinsic change over time. In this connection, Simons points to relatively familiar puzzles concerning the individuation of kinds and their members; identity through time; as well as possible threats of indeterminacy specifically with respect to the boundaries of individuals. In some cases, controversy arises, for example, over whether a given entity (e.g., a coral reef or an aspen grove) really should be regarded as an individual, rather than a collection. In other cases, theorists disagree over how entities, especially those which do not reproduce sexually (e.g., proteins or genes), are best grouped into kinds. Thirdly, the persistence of individuals over time can pose tricky questions for metaphysicians, e.g., how many car-parts one can replace before the old car ceases to exist and a new car has come into being. Finally, everyday substances such as mountains or geographical regions are well-known to be subject to vagueness, as is brought out for example when we try to determine where the Outback begins and ends or whether a particular rock is part of Mt. Everest or the adjacent mountain, Nuptse. What is worse, metaphysical concerns over individuation and persistence, according to Simons' assessment, are not confined to the macroscopic realm, but affect even the micro-physical entities which are referred to and quantified over by our best scientific theories, leading him to speculate pessimistically:

Suddenly it begins to look as though substance, far from being a widely applicable commonplace, is a concept rarely if ever fulfilled, an idealized limit of little or no use to metaphysics. (Ibid., 250)

Based on these considerations, Simons concludes that metaphysicians might as well wake up to the reality that their cherished concept of substance has become obsolete, since in the end there may be nothing at all which meets all seven characteristics listed in (b) and falls under a basic, unanalyzable, primitive concept of substance in compliance with (a). No harm is done, in Simons' view, if we continue to talk about the everyday substances, as long as we are aware of the limitations of this practice. If the concept of substance, as used in the everyday way, retains any usefulness at

all, it does so only as a convenient device for singling out entities that are easily accessible from our specifically human point of view. But there is no reason in general to expect that these anthropomorphically salient objects will retain their special place in light of our most considered scientific and metaphysical theories of the world.

### III. *The continued importance and usefulness of the concept of substance for metaphysics*

#### III.1 *The non-basicness of substance*

Simons' first constraint in (a) states that the concept of substance is to be accepted as a basic unanalyzable primitive. This constraint is open to several interpretations. For one thing, as it was stated above, (a) is most naturally read as concerning the simplicity or complexity of the *concept* of substance itself. Alternatively, Simons may also be targeting positions according to which what is in question is the *metaphysical* simplicity or complexity of the entities to which the concept is applied.<sup>6</sup> Presumably, on either reading, the important question at issue is whether it is possible to explain *in more basic terms* why some particular item is classified as a substance; or whether, instead, once an entity has been designated as a substance, all explanation must be presumed to come to a stop.

Consider in this connection Aristotle's own conception of primary substance in the *Categories*, which forms much of the backdrop for Simons' discussion. Aristotle allows that the entities he classifies there as primary substances are metaphysically complex in a certain way; but the metaphysical complexity in question is not of the right kind to yield a proper explanation as to why some particular item is classified as a primary substance. For, according to the ontology of the *Categories*, entities which are classified as primary substances may have *proper parts* which are themselves

---

6. A third possibility is that Simons takes the simplicity or complexity of the concept of substance to be directly connected to, or even inherited from, the metaphysical simplicity or complexity of the entities to which the concept applies. This third reading is encouraged by Simons' appeal to his own trope-theoretic account of everyday substances, according to which these entities turn out to be complex trope bundles. Simons seems to be suggesting that the concept of substance itself, when used in the everyday way, must be regarded as non-basic as a result of the metaphysical complexity which he attributes to the entities falling under this concept.

classified as primary substances (e.g., Socrates and Socrates' arm).<sup>7</sup> But the mereological complexity Aristotle attributes to these entities does not lead to an explanation in more basic terms as to why some particular item, e.g., Socrates, is classified as a primary substance in the *Categories*. For we could not very well account for Socrates' status as a primary substance in the *Categories* by appeal to the fact that his arms, legs, head, etc., are also classified as primary substances. Presumably, what disqualifies Socrates' proper parts from figuring in an adequate explanation of Socrates' status as a primary substance in the *Categories* is that a mereological analysis of Socrates into his proper parts does not yield an analysis of Socrates into constituents which play a *more basic* explanatory role than Socrates himself, since both Socrates and his proper parts are assigned the *same* status within the ontology of the *Categories*, namely that of primary substances. The occurrence of the term, "basic", in (a) is therefore crucial for a proper understanding of Simons' first constraint, since it is needed to distinguish varieties of complexity which are relevant to the explanatory tasks at hand from those which are irrelevant. (More on these issues below.)<sup>8</sup>

The *Categories* is generally agreed to be one of the earliest, if not *the* earliest, of Aristotle's written works. In contrast, when we turn to such texts as the *Physics*, *De Anima* and the *Metaphysics*, we notice that Aristotle's views concerning metaphysical complexity have undergone a definitive shift. For a variety of reasons, Aristotle comes to believe that such entities as individual living organisms, which were previously among his ontological front-runners in the *Categories*, are further analyzable into *more basic*

---

7. Aristotle is careful to distinguish his "being in a subject" relation from parthood (cf., *Cat.2*, 1a24–25) and is thereby able to maintain in the *Categories* that the parts of substances are themselves substances:

We need not be disturbed by any fear that we may be forced to say that the parts of a substance, being in a subject (the whole substance), are not substances. For when we spoke of things *in a subject* we did not mean things belonging in something as *parts*. (*Cat.5*, 3a29–32) This and all further passages from Aristotle's *Categories* come from the translation by J. L. Ackrill (see Barnes 1984).

8. Given that Simons finds it necessary to appeal to a notion of basicness as well suggests that he himself makes use of some ranking of entities by their degree of explanatory priority. In light of the explanatory roles played by the concept of substance, which will be discussed further below, the resulting ordering of entities by their degree of explanatory priority preserves a key ingredient in the concept of substance, as it is used by metaphysicians. Simons himself therefore does not in the end fully abandon all that is philosophically important in the notion of substance, despite his desire to do so.

constituents, viz., their matter and their form.<sup>9</sup> To illustrate, at the conclusion of his so-called “striptease” argument in *Met. Z.3*, Aristotle assigns a decidedly secondary status to matter/form compounds precisely because of their composite nature:

The substance compounded of both, i.e. of matter and shape, may be dismissed; for it is posterior and its nature is obvious.” (*Met. Z.3*, 1029a30–32; translation by W. D. Ross)<sup>10</sup>

As this passage indicates, although Aristotle still regards matter/form compounds as substances (*ousiai*), these entities have apparently forfeited their status as *primary* substances and are now classified as *posterior*, due to their particular brand of metaphysical complexity. Presumably, once the hylomorphic analysis is on the table, Aristotle has the resources to explain in more basic terms, at least in certain cases, why some particular item is classified as a substance. For example, in response to the question, “But why is Socrates a substance?”, Aristotle can now point to the presence of Socrates’ form (viz., his soul) in his matter (viz., a human body capable of sustaining a human soul). In light of Simons’ first constraint, however, it should strike us as curious that Aristotle would persist in calling matter/form compounds “substances”, even while in the same breath noting that their hylomorphic structure leads them to occupy a posterior explanatory rank relative to their main competitors in his *Metaphysics* ontology, especially form.

On the *conceptual* side as well we find that Aristotle tolerates a certain degree of non-basicality in the notion of substance, even in cases in which the conceptual complexity in question has no direct metaphysical correlate. Famously, Aristotle characterizes the primary substances in the *Categories* as ultimate subjects of predication in the sense that every-

---

9. According to my own mereological reading of Aristotle’s hylomorphism, the matter and the form are construed not only as constituents of a hylomorphic compound but as proper parts, strictly and literally speaking and according to a single notion of parthood (see Koslicki 2008). However, it is not necessary in the current context to enter into the controversial question of whether Aristotle in fact subscribes to the view that the matter and the form are constituents (mereologically or otherwise) of a hylomorphic compound. The important (and hopefully less controversial) point for present purposes is that Aristotle treats matter and form as explanatorily *more basic than* matter/form compounds at least with respect to certain explanatory tasks which take on center-stage in such texts as the *Physics*, *De Anima* and the *Metaphysics*. I will have more to say below concerning the operative notion(s) of explanatory basicness as well as the relevant tasks relative to which explanatory basicness is to be understood.

10. This and all subsequent passages from the *Metaphysics* come from the translation by W. D. Ross (see Barnes 1984).

thing else is either “said of” or “present in” them as subjects (cf., *Cat.5*, 2b3–6).<sup>11</sup> Whether Aristotle intends his characterization in the *Categories* to amount to an outright *definition* of what makes something a primary substance is controversial. But at the very least we can read him as putting forward a *criterion* of some sort which can be construed as either necessary or sufficient for something’s being a substance.<sup>12</sup> Either way, Aristotle’s ultimate subject criterion in the *Categories* provides the basis for an informative answer in conceptually more basic terms to the question of why some particular item deserves to be classified as a primary substance in the *Categories*.<sup>13</sup>

We have in this section encountered reasons for thinking that, at least for Aristotle, the philosophical usefulness and importance of the concept of substance does not hinge on accepting this notion as a basic unanalyzable primitive, as Simons’ first constraint in (a) would suggest. Rather, in Aristotle’s view, the classification of certain entities as substances is in some cases compatible with the availability of an explanation in more basic terms as to why a particular item is classified as a substance. The interesting and difficult task now before us is to elucidate further how the designation of an entity as a substance can contribute to certain central explanatory goals in metaphysics, even when it is admissible that either the concept or the entities themselves are subject to further analysis.

---

11. The first of these relations (“being said of a subject”) is illustrated by an essential predication of the form, “Socrates is human”, in which an individual (Socrates) is said to be a member of a species (human being). The second relation (“being in a subject”) is exemplified in accidental predications of the form, “Socrates is pale”, in which an accidental feature (paleness) is said to inhere in a substantial individual (Socrates).

12. Devereux 2003 for example reads Aristotle’s ultimate subject criterion as providing only a *sufficient* condition for something’s being a substance. Individuals in the category of substance satisfy this criterion by being neither said of nor present in anything else as a subject. In contrast, species and genera in the category of substance fulfil only one half of the ultimate subject criterion: while they are not present in anything as a subject, they are still predicable of other things as subjects. Given that Aristotle in the *Categories* nevertheless classifies substantial species and genera as *secondary* substances, Devereux reasons that the ultimate subject criterion can only be read as providing a sufficient, but not also a necessary, condition for substancehood. There is, however, also the possibility of interpreting the ultimate subject criterion as a condition that is both necessary and sufficient for an entity’s status as a *primary* substance in the *Categories*.

13. The complexity at issue in the *Categories* is merely conceptual and lacks a direct metaphysical correlate, since entities which are classified as primary substances in the *Categories* are *not* taken to be *composed* of an ultimate subject of predication as a constituent, along with other constituents. Rather, Aristotle in the *Categories* thinks of these entities as themselves *being* (identical to) ultimate subjects of predication.

### 3.2 *Taxonomic vs. non-taxonomic substancehood*

Among the most important explanatory roles played by the concept of substance in philosophical contexts are what I will call a “*taxonomic*” and a “*non-taxonomic*” role. First, in its taxonomic role, philosophers employ this notion to single out certain *kinds* of entities (e.g., macroscopic concrete particular objects), without thereby simultaneously committing themselves to the idea that these entities must be assigned a privileged ontological position within their respective ontologies. When the concept of substance is utilized in this first taxonomic way, the substances appear merely as one among many entries in a catalogue of beings. The resulting inventory might for example constitute an answer to the existential question, “What is there?”, which Quine saw as central to the discipline of ontology. In other contexts, however, philosophers employ the concept of substance in a second non-taxonomic role, in order to indicate that certain kinds of entities (taxonomically speaking) deserve to be singled out for special treatment in the ontology in question. Much confusion has resulted over the years from a failure to distinguish between these two very different, but equally important, roles played by the concept of substance in philosophical contexts.

To illustrate, consider Aristotle’s well-known list of ten categories in which “substance” appears as his first entry:

Of things said without any combination, each signifies either substance or quantity or qualification or a relative or where or when or being-in-a-position or having or doing or being-affected. (*Cat.4*, 1b25–27)

Here, Aristotle draws our attention to a ten-fold division among different *kinds* of being (taxonomically speaking). In order to set up a hierarchical ordering among these entities, Aristotle appeals to his two relations, “being in a subject” and “being said of a subject”, as follows:

A *substance*—that which is called a substance most strictly, primarily, and most of all—is that which is neither said of a subject nor in a subject, e.g. the individual [human being] or the individual horse. The species in which the things primarily called substances are, are called *secondary substances*, as also are the genera of these species. For example, the individual [human being] belongs in a species, [human being], and animal is a genus of the species; so these—both [human being] and animal—are called secondary substances. (*Cat.5*, 2a11–19; Ackrill’s italics)<sup>14</sup>

14. I have taken the liberty of replacing “man” with “human being” in Ackrill’s rendition of this passage.

Given his ultimate subject criterion, certain kinds of entities (taxonomically speaking), e.g., individual living organisms, turn out to be “most strictly, primarily, and most of all” deserving of the title, “substance”, in Aristotle’s *Categories* ontology, while others rank below them: the so-called “secondary substances” (viz., the species and genera to which substantial individuals belong); as well as the *non*-substances (viz., the individuals, species and genera belonging to the other nine taxonomic categories listed earlier).

In view of these importantly different roles played by the notion of substance in philosophical contexts, it now emerges that Simons’ everyday and metaphysical uses of the concept of substance, distinguished earlier in (i) and (ii), work at cross purposes. For when Simons speaks of the *everyday* substances, he is most naturally read as having in mind a *taxonomic* division among entities. In this context, it would be decidedly odd to ask for example, “Is Socrates an *everyday* substance?”, since it is taken as fixed that individual living organisms count as paradigm everyday substances.<sup>15</sup> In contrast, when Simons speaks of the *metaphysical* substances, he has in mind a *non-taxonomic* use of the concept of substance. In this context, it is perfectly legitimate to ask, “Is Socrates a *metaphysical* substance?”. In fact, individual living organisms, even though they count as paradigmatic everyday substances in the context of Simons’ discussion, turn out not to be metaphysical substances, given the constraints in (a) and (b).

The concept of substance, then, can be used by metaphysicians, among other things, in the service of two distinct explanatory tasks: first, to classify entities taxonomically; and, secondly, to impose a non-taxonomic ordering of some kind onto the entities included in a given ontology. The second project moreover requires metaphysicians to engage with the further question of *why* certain items are to be assigned a privileged position within a particular ontology. For this reason, philosophical disputes concerning substancehood often center on the very *criteria* themselves by means of which substance status is awarded to certain taxonomic categories within a given ontology. Simons, for the purposes of his discussion, settles on a particular choice of criteria; but not everyone will agree with Simons’ preference. (More on these issues below in Section III.4.)

---

15. Even in its taxonomic uses, however, the concept of substance may admit of interesting borderline cases. In a context in which it is presupposed that individual living organisms are paradigmatic everyday substances, for example, one may nevertheless wonder about the status of bacteria, if bacteria constitute a borderline case of living organisms.

### 3.3 *Absolute, comparative and relational substancehood*

The concept of substance in its philosophical applications can take on further explanatory roles depending on whether it is used in an *absolute*, a *relational*, or a *comparative* way:

ABSOLUTE SUBSTANCEHOOD: x is a substance *simpliciter*.

RELATIONAL SUBSTANCEHOOD: x is *the substance of* (or *a substance of*) y.

COMPARATIVE SUBSTANCEHOOD: x is *more deserving of substance status than* y.

In its first absolute role, the concept of substance is used to designate entities as substances *simpliciter*. In its second relational role, the concept of substance picks out a relation between pairs of entities, x and y, when x is *the substance of*, or *a substance of*, y. When used in the third comparative way, the concept of substance ranks entities by the degree to which they are *deserving of substance status*.

All three of these explanatory roles are evident in Aristotle, but unfortunately not always clearly distinguished by Aristotle himself or his commentators. For example, at *Cat.*5, 2a11–19 (cited earlier), Aristotle classifies individual living organisms, as well as their species and genera, as substances *simpliciter*, using an absolute notion of substancehood. In the very same passage, however, we find Aristotle identifying certain items as *primary* substances (e.g., individual living organisms) and others as *secondary* substances (e.g., the species and genera to which these individual living organisms belong). In this usage, Aristotle ranks entities in a comparative way, by the degree to which they satisfy the ultimate subject criterion at work in the *Categories*.<sup>16, 17</sup> The third relational use of the concept of substance is operative for example in the opening lines of *Met.*Z.6:

We must inquire whether each thing and its essence are the same or different. This is of some use for the inquiry concerning substance; for each thing is thought to be not different from *its* substance [*tês heautou ousia*], and the

---

16. See also for example *Cat.*5, 2b7–8: “Of the secondary substances the species is *more a substance than* [*mallon ousia*] the genus, since it is nearer to the primary substance”.

17. Aristotle also holds that no species is any more a secondary substance than any other species; and no individual is any more a primary substance than any other individual (cf., *Cat.*5, 2b22–24). This statement is of course compatible with the idea that, on the whole, the primary substances are more deserving of substance status (i.e., more substances) than the secondary substances.

essence is said to be *the substance of* each thing [*hē hekastou ousia*].” (*Met.Z.6, 1031a15–18*; my italics)

In this passage, Aristotle speaks of the essence of each thing, relationally, as *the substance of* that thing.<sup>18</sup>

Simons, in his discussion, appears to be exclusively concerned with an absolute conception of substance, according to which entities are classified as either everyday substances or metaphysical substances *simpliciter*. Much of the philosophical usefulness and importance of the notion of substance, however, stems not only from its applications as an absolute concept, but also from its relational and comparative uses. When an entity (e.g., Socrates’ form) is designated relationally as *the substance of*, or *a substance of*, another (e.g., Socrates), we can expect the first entity to answer certain specifically metaphysical questions which arise in connection with the second entity, e.g., questions concerning existence, identity, essence, parthood, dependence, unity, and the like. (See Section III.5 for further discussion.) Moreover, as we have observed in earlier sections, when an entity is classified in a comparative way as *more deserving of substance status than* (or *more of a substance than*) another, the resulting ordering of entities indicates the degree to which a certain taxonomic category of entities is assigned a privileged position in a given ontology.

---

18. It is exceedingly difficult to determine how exactly Aristotle conceives of the connections between his absolute, relational and comparative uses of the concept of substance. It is by no means clear that Aristotle wants to be committed to the following principle connecting relational and absolute uses of the concept of substance: if *x* is *the substance of* (or *a substance of*) *y*, then *x* is a substance *simpliciter*. In fact, in a case in which *y* is a substance *simpliciter* and *x* is *actually* present in *y*, this inference would lead to trouble with a principle Aristotle endorses in *Met.Z.13*, according to which no substance can have other substances present in it actually (cf., 1039a2–14). (The *Met.Z.13* principle would contradict Aristotle’s earlier view from the *Categories*, discussed in Section III.1, according to which a primary substance can have proper parts which are themselves primary substances, but only under the assumption that the proper parts of a primary substance are present in it actually.) Moreover, consider the following attempt at connecting comparative and absolute uses of the concept of substance: *x* is a substance *simpliciter* just in case *x* is *more deserving of substance status* than any other entity relative to *all* comparative criteria of substancehood. Such a principle has the consequence that nothing would be classified as a substance *simpliciter*, since no single type of entity meets *all* the comparative criteria of substancehood Aristotle invokes along the way. Unfortunately, I will have to leave the further investigation of these interesting questions for another occasion.

### 3.4 Preferred cases and criteria

In his most pessimistic moments, Simons speculates that perhaps nothing at all qualifies as a metaphysical substance. Given that he thinks of metaphysical substancehood in an absolute non-taxonomic way, Simons' assessment does not directly speak to the philosophical usefulness and importance of the concept of substance in its other explanatory roles. Moreover, whether nothing at all occupies the ontologically privileged position associated with absolute non-taxonomic uses of the concept of substance of course depends crucially on the *criteria* of substancehood which are invoked in a particular context. Simons operates with a conception of metaphysical substancehood according to which this notion is governed by the constraints specified in (a) and (b). Under this construal, he arrives at the conclusion that the concept is only ever satisfied (if at all) under very special circumstances, and not generally by the macroscopic concrete particular objects which figure prominently in our human experience of the world or by the entities referred to and quantified over by our best scientific theories. But Simons' strategy raises the question of why those who are sympathetic to an absolute non-taxonomic conception of substancehood should accept these particular constraints as reflective of what they have in mind when they assign a privileged ontological position to certain taxonomic categories of entities.

Consider for example Simons' own ontology, according to which everyday substances turn out to be complex trope bundles and are hence classified as non-basic, both on conceptual and metaphysical grounds. We may nevertheless wonder whether this ontology in fact contains metaphysical substances after all, only (contrary to what we might have expected) the metaphysical substances, in this ontology, would turn out to be *tropes*, rather than everyday substances. Under this construal, Simons' first constraint in (a) would presumably be satisfied, since he takes tropes to be both conceptually and metaphysically basic. But we would nevertheless run into difficulties with respect to the second constraint in (b), since even tropes, on Simons' conception, do not meet all seven of the characteristics cited above. Among other things, tropes turn out not to be ontologically independent, as some of their existential needs can only be met when they find themselves in the presence of other tropes in the form of certain kinds of trope bundles. Furthermore, given that tropes are particularized property instances, they do not function as ultimate subjects of predication; rather, they are what is predicated of a subject, viz., the

trope bundle. (See Simons 1994 for a defense of this particular version of trope theory.)

At this point, however, we face a decision: we can either retain (a) and (b) and continue to operate with a substance concept governed by these constraints; or we can reject (a) or (b) and opt instead for an alternative conception of substancehood that is governed by a different set of constraints. The first route, if Simons is right, leads to an outmoded substance concept with little or no application to the real world, as presented to us by commonsense or science. The second route, however, is very much a live option as well for those who are convinced that the concept of substance retains its usefulness and importance for metaphysics. In the past, philosophical disputes concerning substancehood have often focused precisely on the very *criteria* themselves by means of which certain taxonomic categories of entities are assigned a special role relative to a particular ontology. Such philosophical disputes concerning the criteria of substancehood have in no way been closed off by the considerations Simons brings to bear on this discussion.

### 3.5 *Metaphysical explanations*

Armed with these distinctions, we can now return to the puzzle I raised earlier in connection with Aristotle's pronouncement at the end of his "striptease" argument in *Met.Z.3*, repeated here:

The substance compounded of both, i.e. of matter and shape, may be dismissed; for it is posterior and its nature is obvious. (*Met. Z.3*, 1029a30–32)

Given that Aristotle classifies matter/form compounds, here and elsewhere, as substances *simpliciter*, using an absolute (rather than a relational or comparative) construal of the notion of substance, we can further refine the question at issue and ask why Aristotle is willing to designate an entity as a substance *simpliciter*, despite assigning to it an explanatorily posterior rank.

The apparent oddity inherent in Aristotle's practice is alleviated by taking substancehood *simpliciter* in this context as playing a *taxonomic*, rather than a non-taxonomic, role. In this way, the occurrence of "ousia" in the above passage from *Met.Z.3* is comparable to that in Aristotle's list of the ten categories cited earlier (cf., *Cat.4*, 1b25–27). As the *Met.Z.3* passage brings out, it would be a mistake to assume in general that whatever is classified as a substance *simpliciter*, in a taxonomic sense, will also be desig-

nated as explanatorily prior relative to a non-taxonomic use of the notion of substancehood. In fact, Aristotle attempts to establish in *Met.Z* that form is *more deserving of substance status than* both matter and the matter/form compound with respect to various non-taxonomic comparative considerations he invokes along the way. And yet, ironically, Aristotle is also hesitant to classify forms as substances *simpliciter*.<sup>19</sup> Instead, he usually opts for the relational designation of form as *the substance of* the matter/form compound.<sup>20</sup>

Relative to this ontology, then, we encounter the following combination of classifications:

- (1) A matter/form compound is a substance *simpliciter*.
- (2) A matter/form compound is *less deserving of substance status* than its form.
- (3) A form is *the substance of* a matter/form compound.
- (4) A form is *not* a substance *simpliciter*.

Given the different explanatory roles played by the notion of substance, it makes perfect sense from Aristotle's point of view to designate matter/form compounds as substances *simpliciter* (using a taxonomic absolute conception of substancehood), despite the fact that these entities are also ranked as *less deserving of substance status*, and hence as explanatorily posterior, than their forms (using non-taxonomic comparative criteria of substancehood), while form is in turn classified only relationally as *the substance of* a matter/form compound.

After Aristotle announces, in the passage from *Met.Z.3* cited above, that the substance composed of matter and form may be "dismissed" on the grounds that "it is posterior and its nature is obvious", he continues as follows:

It is agreed that there are some substances among sensible things, so that we must look first among these. For it is in an advantage to advance to that which

---

19. Aristotle's qualms about designating forms as substances *simpliciter* might at least in part be traced to his adherence to the *Met.Z.13* principle, mentioned earlier, according to which no substance (*simpliciter*) can be present in another substance (*simpliciter*) actually. Aristotle seems to be motivated in this connection primarily by concerns over unity.

20. See for example the closing paragraph of *Met.Z.17*. Although Aristotle here does not explicitly mention form, he is usually read as having in mind form when he speaks of *the substance of* a matter/form compound, which he also characterizes as its principle, its nature and the primary cause of its being (cf., 1041b27–29).

is more intelligible. For learning proceeds for all in this way –through that which is less intelligible by nature to that which is more intelligible; and just as in conduct our work is to start from what is good for each and make what is good in itself good for each, so it is our work to start from what is more intelligible to oneself and make what is intelligible by nature intelligible to oneself.” (*Met.Z.3*, 1029a33–1029b12)

Like everyone else, then, metaphysicians start with what is salient to them and ask questions about these entities which are specific to their discipline, e.g., questions concerning existence, identity, essence, parthood, dependence, unity, and the like. The entities which form the starting-point, but by no means the end-point, of metaphysical inquiry are called “substances” in a taxonomic sense (*viz.*, Simons’ everyday substances). In the course of their investigation, metaphysicians may well find themselves “dissolving” these experientially salient objects into explanatorily more basic constituents: e.g., matter and form (in Aristotle’s case) or tropes (in Simons’ case). But progress in metaphysics, *i.e.*, advancing from “what is intelligible to us” to “what is intelligible by nature”, does not require that anything in the end be classified as a substance *simpliciter* in an absolute non-taxonomic sense (Simons’ metaphysical substances). Rather, we may well end up with an ontology in which what is highlighted as explanatorily prior qualifies for substance status only in a relational or comparative sense. All the while, we are nevertheless well within the bounds of the study of being *qua* being as focused on the principles and causes of substances:

But everywhere science deals chiefly with that which is primary, and on which the other things depend, and in virtue of which they get their names. If, then, this is substance, it will be of substances that the philosophers must grasp the principles and the causes. (*Met. Γ.2*, 1003b16–19)

#### 4. Conclusion

In this paper, I have argued that the concept of substance retains its importance and usefulness for metaphysics, despite Simons’ powerful considerations to the contrary. First, the philosophical significance of this notion does not require that either the concept itself or the entities falling under it are accepted as basic unanalyzable primitives. Secondly, philosophers employ the concept of substance not only to differentiate among different taxonomic categories of entities (e.g., Simons’ everyday substances), but also to mark some of these as occupying a privileged

ontological position (e.g., Simons' metaphysical substances). Thirdly, philosophically significant uses of the concept of substance for metaphysics do not necessitate that anything be classified as a substance *simpliciter* in an absolute non-taxonomic sense (Simons' metaphysical substances), since the concept can also be employed relationally and comparatively ("x is *the substance of* or *a substance of* y"; "x is *more deserving of substance status than* y"). Fourth, whether anything does qualify for substance status in one of these senses crucially depends on the criteria of substancehood themselves which are at play in a particular context. Simons proceeds by selecting one particular set of criteria; but any such choice is of course controversial.

Are the everyday substances metaphysical substances? Are there any metaphysical substances at all? How these questions are answered depends on the particular use of the concept of substance at issue and the constraints governing this use. Whether any particular such use is to be preferred over others itself requires a philosophical discussion involving the concept of substance. Metaphysicians therefore ought to hold on to their concept of substance; it has by no means outlived its usefulness and importance for their discipline.<sup>21</sup>

#### REFERENCES

- Barnes, Jonathan (ed.) 1984: *The Complete Works of Aristotle: The Revised Oxford Translation*. Volume One and Two, Bollingen Series LXXI.2. Princeton: Princeton University Press.
- Corkum, Phil 2008: "Aristotle on Ontological Dependence". *Phronesis* 53, 65–92.

---

21. Material from this paper was presented in a seminar I taught at the University of Alberta on theories of fundamentality during the Winter of 2014; as well as at talks at the University of Victoria and Concordia University. I thank those present at these events for their extremely helpful feedback. In particular, I greatly benefited from conversations with Margaret Cameron, Phil Corkum, Klaus Jahn, Colin Macleod, Michael Raven, Esther Rosario, Andrew Tedder, Lianghua Zhou, Jack Zupko and Justin Zylstra. In addition, my thoughts on Aristotle's conception of substance were very much shaped by a seminar I taught on *Metaphysics* Z, H, Θ at the University of Colorado-Boulder during the Fall of 2013. I would like to acknowledge the insightful contributions of those who participated in this seminar, in particular Philip Choi, Daniel Coren, Erlantz Etxeberria, Jay Geyer, Mitzi Lee, Bob Pasnau, Caleb Pickard, Chaz Vollmer, Joe Wilson and Alex Wolf-Root.

- Corkum, Phil 2013a: "Substance and Independence in Aristotle". In: Miguel Hoeltje, Benjamin Schnieder & Alex Steinberg (eds.), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence*. Munich: Philosophia Verlag, 65–95.
- 2013b: "Critical Notice for Michail Peramatzis's *Priority in Aristotle's Metaphysics*". *Canadian Journal of Philosophy* 43, 136–156.
- Devereux, Daniel 2003: "The Relationship between Books Zeta and Eta of Aristotle's *Metaphysics*". *Oxford Studies in Ancient Philosophy* 25, 159–211.
- Fine, Kit 1995: "Ontological Dependence". *Proceedings of the Aristotelian Society* 95, 269–290.
- Gorman, Michael 2005: "The Essential and the Accidental". *Ratio (New Series)* 18, 276–289.
- 2006a: "Substance and Identity-Dependence". *Philosophical Papers* 35, 103–118.
- 2006b: "Independence and Substance". *International Philosophical Quarterly* 46, 147–159.
- Gorman, Michael & Sanford, Jonathan (eds.) 2004: *Categories: Historical and Systematic Essays*. Studies in Philosophy and the History of Philosophy 14. Washington: The Catholic University of America Press.
- Hoffman, Joshua & Rosenkrantz, Gary 1994: *Substance Among Other Categories*. New York: Cambridge University Press.
- Hoffman, Joshua & Rosenkrantz, Gary 1997: *Substance: Its Nature and Existence*. London: Routledge.
- Koslicki, Kathrin 2008: *The Structure of Objects*. Oxford: Oxford University Press.
- 2013a: "Substance, Independence and Unity". In: Edward Feser (ed.), *Aristotle on Method and Metaphysics*. Basingstoke: Palgrave/Macmillan, 169–195.
- 2013b: "Ontological Dependence: An Opinionated Survey". In: Miguel Hoeltje, Benjamin Schnieder & Alex Steinberg (eds.), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence*. Munich: Philosophia Verlag, 31–64.
- 2015: "The Coarse-Grainedness of Grounding". *Oxford Studies in Metaphysics* 9, 306–344.
- Lowe, E. J. 1998: *The Possibility of Metaphysics: Substance, Identity, and Time*. Oxford: Oxford University Press.
- 2006: *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Clarendon Press.
- Peramatzis, Michail 2008: "Aristotle's Notion of Priority in Nature and Substance". *Oxford Studies in Ancient Philosophy* 35, 187–247.
- 2011: *Priority in Aristotle's Metaphysics*. New York: Oxford University Press.

- Schaffer, Jonathan 2009: "On What Grounds What". In: David Chalmers, David Manley & Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Clarendon Press, 347–383.
- Simons, Peter 1987: *Parts: A Study in Ontology*. Oxford: Oxford University Press.
- 1994: "Particulars in Particular Clothing: Three Trope Theories of Substance". *Philosophy and Phenomenological Research* 54, 553–575.
- 1998: "Farewell to Substance: A Differentiated Leave-Taking". *Ratio (New Series)* 11, 235–252.
- Toner, Patrick 2008: "Emergent Substance". *Philosophical Studies* 141, 281–297.
- 2010: "On Substance". *American Catholic Philosophical Quarterly* 84, 25–48.
- 2011: "Independence Accounts of Substance and Substantial Parts". *Philosophical Studies* 155, 37–43.

HOW TO SPEAK OF EXISTENCE:  
A BRENTANIAN APPROACH TO (LINGUISTIC AND  
MENTAL) ONTOLOGICAL COMMITMENT

Uriah KRIEGEL  
Institut Jean Nicod

*Summary*

To a first approximation, ontology is concerned with what exists, metaontology with what it means to say that something exists. So understood, metaontology has been dominated by three views: (i) existence as a substantive first-order property that some things have and some do not, (ii) existence as a formal first-order property that everything has, and (iii) existence as a second-order property of existents' distinctive properties. Each of these faces well-documented difficulties. In this chapter, I want to expound a fourth theoretical option, which unfortunately has remained 'under the radar.' This is Franz Brentano's view, according to which to say that X exists is not to attribute a property at all (first- or second-order), but to say that the correct attitude to take toward X is that of accepting or believing in it.

1. *Introduction: Metaontology and existence talk*

Moral philosophy is usefully divided into ethics and metaethics. Oversimplifying considerably, the distinction is this: ethics is concerned with which things are good, metaethics with what it means to say that something is good. The goal of ethics is to produce a comprehensive list of all the good things (in the broadest sense of the term). Metaethics concerns a more fundamental question: when we say that X is good, what exactly are we saying? In a way, ethics is concerned with the *extension* of the concept GOOD, metaethics with its *intension*.

This is an oversimplification in at least two ways. First, ethics and metaethics are concerned with other normative concepts, such as RIGHT, VIRTUE, and REASONS. Secondly, metaethics deals with other issues, such as moral epistemology—how we can come to know what things are good. Still, there

is a clear sense in which answering the question of what exactly we are doing when we say that something is good lies at the heart of metaethics.

A similar division of labor may be applied to ontology and metaontology. Again oversimplifying, ontology is concerned with what exists, metaontology with what it means to say that something exists. The goal of ontology is to produce a comprehensive list of existents; that of metaontology is to answer the question of what exactly we are saying when we say that X exists. To that extent, ontology is concerned with the extension of the concept EXISTENCE, metaontology with its intension.

One way in which this oversimplifies is that ontology may well be concerned with other concepts, such as GROUNDING, FUNDAMENTALITY, or ESSENCE.<sup>1</sup> Another is that metaontology is also concerned with other issues, notably the methodology of ontology.<sup>2</sup> Nonetheless, there is a sense in which at the heart of metaontology lies the question ‘when we say that X exists, what exactly are we saying?’

To this question, there are three prominent answers in the extant literature. According to the first, to say that X exists is to attribute to X a substantive, discriminating first-order property that some things have and some do not (Meinong 1904, Parsons 1980). According to the second, it is to attribute a second-order property of existents’ distinctive properties or of the concept designed to pick them out (Frege 1884, Russell 1905). According to the third answer, more popular in recent discussions, to say that X exists is to attribute to X a formal, *undiscriminating* first-order property that everything has (Williamson 2002, van Inwagen 2003). Each of these has met with strong resistance and faces extraordinary objections, but have also been admirably defended. My main goal here is to present a fourth alternative, drawn from Brentano’s metaontology. Just by way of motivating the search for another approach to existence talk, §2 offers a brief survey of these views and some of their immediate difficulties.

---

1. See Schaffer (2009) for a view of ontology as concerned primarily with grounding and fundamentality rather than existence, and Lowe (2008) for the view that essence is a central part of what ontology is about.

2. Thus, debates over Quine’s (1948) quantificational method vs. Armstrong’s (2004) truth-maker method belong within the sphere of metaontology.

## 2. *Three approaches to existence claims*

The simplest view is that to say that X exists is to attribute a substantive property to X. When I say that Obama is president, I attribute to Obama a certain property, namely the property of being president. In exactly the same manner, when I say that Obama exists, I attribute to him a property, this time the property of existing. President, existent, confident, American—those are all Obama-esque attributes on a par. Accordingly, existential claims are at bottom of a kind with predicative claims: ‘ducks are cute’ and ‘there are ducks’ *look* different, but the latter is just an unhelpful rendering of ‘ducks are existent.’

Dissatisfaction with this approach is rife. There are technical problems to do with negative existentials and existential generalization. From ‘Jimmy is not president’ I can infer ‘there is a non-president.’ If existential claims work just like predicative ones, from ‘Shrek does not exist’ I should be able to infer ‘there is a nonexistent.’ But this requires a distinction between ‘there is’ and ‘exists’ that many find odious. There are also non-technical problems: as Hume (1739 I,II,vi) noted, the idea of existence adds nothing to the idea of an object. The idea of a cute duck is different from the idea of a duck, which means that the idea of cuteness contributes something to the idea of a cute duck. But the idea of an existing duck is nowise different from the idea of a duck; so it is unclear what the idea of existence is supposed to contribute.

Perhaps the most dominant view historically is that in saying that X exists we are attributing a property not to X, but either (i) to X’s distinctive, individuating *properties* or (ii) to the *concept* of X. In the first case, we attribute the property of being instantiated; in the second, that of referring. In both versions, existence is construed as a *second-order* property, since it is not a property of X itself but of some properties of X or the concept of X. Thus when I say that Obama exists, what I am doing is attributing to the properties that individuate Obama (whatever they are) the property of being instantiated, or else attributing to the concept OBAMA the property of referring. Likewise, when I say that dragons do not exist, I am saying that dragon-hood is uninstantiated, or else that DRAGON is empty.

This approach raises its own set of difficulties. Some are technical and pertain to its application to singular existentials. The approach can be applied to ‘Obama exists’ only if the proper name ‘Obama’ is semantically associated with certain properties. For example, if ‘Obama’ just means ‘the actual 44<sup>th</sup> US president,’ then perhaps in saying that Obama exists we are

saying that the property of being the actual 44<sup>th</sup> US president is instantiated (or else that the concept of the actual 44<sup>th</sup> US president refers). But many philosophers deny that ‘Obama’ is associated with *any* properties, holding instead that it refers *directly* to the individual himself, without mediation by properties (or concepts) (Kripke 1972). There are also non-technical problems with the approach: it implies that in saying that Obama exists, we are not saying anything *about Obama*; in fact, we are not speaking of Obama at all, but of some different entity. What we are speaking of is not even a concrete particular, but a property or a concept. This feels wrong: saying that X exists feels like a comment on X, not on something else suitably related to X.<sup>3</sup> When we exclaim excitedly that the Higgs boson exists, it is the discovery of the boson that excites us.

A view gathering momentum in recent metaontology is that existence is a first-order property of things, but not a substantive, discriminating one that divides entities into two subsets, those which have the property and those which do not. Rather, it is a formal or pleonastic property that everything has. Other logical or formal properties are like that as well: the property of being self-identical does not divide entities into two subsets either.

Since the view is more recent, there have not yet emerged standard objections to it in the literature. But one immediate worry is that it is unclear on this view how we might *acquire the concept* of existence. The most basic way to acquire the concept of F by interacting sufficiently with Fs and non-Fs to develop a sensitivity to the difference between them. But if existence were a formal property of everything, this kind of differential interaction with existents and nonexistents would be ruled out. Some concepts we acquire purely by putting together other concepts acquired through differential interaction: we can acquire the concept BACHELOR, for example, by putting together MAN and UNMARRIED. However, we can do this only where there are *genus et differentia*. If existence is a property of *everything*, it cannot be a species of any other, more generic property. Finally, it might be claimed that EXISTENCE is simply an *innate* concept. But for the view not to be mysterian, it must construe innate concepts as acquired phylogenetically rather than ontogenetically through the same two mechanisms that we have just ruled out for existence as a universal property.

---

3. See Frege (1884, 67) for the explicit claim that ‘X exists’ is not about X, and Thomasson (2015 Chap. 2) for criticism of it.

Some philosophers have proposed to adopt all three notions of existence but apply them to different regions of our existence discourse (Voltolini 2012). For example, one might hold that certain simples arranged table-wise exist in the sense that they exhibit the substantive first-order property, the table they compose exists in the sense that it exhibits the formal first-order property, and the artifactual kind Table (to which it belongs) exists in the sense that it exhibits the second-order property of being instantiated. Note, however, that the mentioned problems of the three views of existence are not problems for the claim that the relevant view accounts for *all* cases of existence, but for the claim that it accounts for *any*. Therefore, in committing to all three kinds of existence, one would be *amassing* all three views' problems, rather than avoiding any of them.

To be sure, proponents of each view have offered various responses to these and other problems facing them. I do not wish to dwell on these matters here. Instead, I want to articulate an alternative approach—a coherent and stable account of existence talk that has not as yet received a proper airing. This fourth alternative was developed by Franz Brentano (1930, 1933), but has seen virtually no uptake outside the circles of Brentano scholarship.<sup>4</sup> My goal is to motivate Brentano's position to a wider audience and show that it merits serious consideration. I start by drawing out three assumptions shared by the three more familiar views of existence talk; at least two of them are rejected in the Brentanian alternative.

### 3. *Commitment to existence: Linguistic and mental*

To say that X exists is to perform a certain linguistic act. The performance of this act commits the performer to X's existence. To that extent, we may think of the act of saying that X exists as *linguistic existence-commitment*. Asserting 'X exists' is of course only one form of linguistic existence-commitment. Other include asserting 'there is an X,' 'there exist no Xs,' 'X is,' 'the Xs are existent,' and so on.

It is, of course, possible to commit oneself to the existence of X without saying anything. I may think to myself that X exists and keep the thought to myself. This would also be a form of existence-commitment,

---

4. Brentano worked on this in two main periods of his life. His doctoral dissertation was on the notion of existence in Aristotle (Brentano 1862), but he returned to the topic forty years later and composed (or dictated, once he turned blind circa 1907) a number of important essays and lecture notes; most are collected in Brentano 1930, some in Brentano 1933.

but not of *linguistic* existence-commitment. Rather, it is a form of *mental* existence-commitment. Mental existence-commitment is commitment to something's existence *in thought*, whereas linguistic existence-commitment is commitment to something's existence *in language*.

One assumption shared by all three familiar views of existence talk is that linguistic existence-commitment is a matter of attributing a property to something. The three views differ on what property is attributed and what it is attributed to, but they agree that some property is attributed to something.

A second shared assumption is that mental existence-commitment, like linguistic existence-commitment, is a matter of attributing a property to something, indeed the same property to the same thing. On one view to judge that Obama exists is to mentally attribute a substantive first-order property to Obama, on another it is to mentally attribute a formal first-order property to Obama, and on a third view it is to attribute a second-order property to the Obamarific properties (or to OBAMA).

The third shared assumption follows straightforwardly from the other two. It is that there is no difference between linguistic and mental existence-commitment other than that one is linguistic and the other mental. The two are structurally the same, but carried out in different representational media. There is no *deep* difference in the mechanics of commitment to existence, it is just that one kind of commitment is linguistically encoded while the other is mentally encoded.

All three assumptions are very natural to make, but as we will see, Brentano rejects the first two (and even the third looks very different in his account). Like many modern philosophers of mind, Brentano presupposes the priority of the mental over the linguistic.<sup>5</sup> Accordingly, he starts from an account of mental existence-commitment, on which basis he devises an account of linguistic existence-commitment. These are taken up in §§4–5.

#### 4. *Mental existence-commitment: Brentano's attitudinal account*

When I think to myself that Obama exists, I mentally commit to the existence of Obama. As noted, the three familiar views share the assumption

---

5. This notion can take several forms, but prominently, it is a widespread view that linguistic representation derives from mental representation, and accordingly that linguistic *content* (meaning) derives from mental content (Grice 1969, Searle 1983, Cummins 1989). All these claims seem to be presupposed by Brentano.

that in doing so, I attribute a property to something. Underlying this is an even deeper assumption: that the commitment to Obama's existence is an aspect of the relevant thought's *content*. The property attributed is a constituent of the content of my thought. On the first-order views, the content is <Existence, Obama>; on the second-order one, it is <Instantiatedness, Obamarific properties>. Either way, some existence-related property figures in the content of existence-committing mental acts.

Brentano rejects this too. For him, mental commitment to something's existence is not an aspect of a thought's content, but of its attitude. The content/attitude terminology is modern: we say that a belief that *p* and a belief that *q* involve the same attitude toward different contents, but that a belief that *p* and a desire that *p* involve different attitudes toward the same content.<sup>6</sup> Put in those terms, the Brentanian idea is that existence-commitment is an *attitudinal property* of some mental states. A mental state that commits to the existence of *X* is not one that represents *X*'s existence, but one that simply represents *X*, but does so in an existence-affirming *manner*. The existence-affirmation is not a dimension of *what* the state represents but of *how* it represents.

We might put the point as follows: mental commitment to *X*'s existence is not a matter of representing *X* as existent, but a matter of representing-as-existent *X*. This formulation is intended to bring out that the existential element is a modification of the representing, not a part of the represented. On this view, to think that Obama exists is to represent-as-existent Obama. The content of the thought is thus exhausted by Obama. Existence does not come into the thought at the level of content, but at the level of attitude. Some of our attitudes exhibit this attitudinal feature, some do not. All and only those that do incorporate a commitment to the existence of what shows up in their content. But the commitment itself does not show up in the content. It is an aspect of the attitude exclusively. Call this the *attitudinal account of mental existence-commitment*.

What motivates the attitudinal account for Brentano is reflection on the difference between believing that *X* exists and merely contemplating that *X* exists. The content of these two acts is identical, but one embodies mental commitment to *X*'s existence while the other does not. Since any content that can be believed can also be contemplated, but contemplation never

---

6. Brentano's terminology is different: his contrast is between the 'object' and the 'mode' of intentionality. To love a mockingbird is to represent an object, the mockingbird, under a mode, the love-mode. Brentano's claim is that commitment to Obama's existence is an aspect of an intentional state's mode, not object.

commits to the reality of what is contemplated, existence-commitment cannot come from the belief's content.<sup>7</sup>

(The point can be put more impressionistically in terms of the difference between fiction and non-fiction prose. In the opening pages of his novel *Father Goriot*, Balzac writes that what he is about to tell us is not just a story—'all is true.' But of course that is just part of the story. Nothing Balzac can write inside his story can make it more than a story. For his novel's status as fiction is not determined by anything inside the novel. The *cover* of a book can announce that it is fiction, or that it is non-fiction, but nothing in the book's pages can confine it to, or rescue it from, the status imputed on it by the cover. Analogously, nothing in the content of a mental act can determine whether what the act represents is intended as real or unreal. The committal or non-committal status of the act must come from outside its content, from the attitude it employs.)

If the existence part of my mental commitment to Obama's existence does not show up in the content of my thought, and the thought's content is exhausted by Obama, then we are dealing here with an *objectual* rather than *propositional* attitude, akin to fear and love.<sup>8</sup> It might seem odd to posit a *cognitive* attitude directed at objects and not propositions or states of affairs. However, we do speak not only of belief-that but also of belief-*in* (as in 'Jimmy believes in ghosts'). Belief-*in* is clearly a cognitive objectual attitude: the content of Jimmy's state is exhausted by ghosts, while the commitment to their existence comes from the attitude of believing-*in*.<sup>9</sup> The basic idea, then, is that to mentally commit to the existence of Obama is not to mentally attribute any property to Obama or any associated entity, but simply to adopt the attitude of believing-*in* toward Obama. Indeed, we may say that believing in Obama is a mental state whose content is Obama and whose attitude is characterized by the property of representing-as-existent.

---

7. Consider: "There have been some psychologists who have maintained that the belief in an object, the affirmation of it, consists in a compounding of presentations. For example, in the judgment "a tree exists," I would affirm the tree as subject and would add as predicate the presentation of something existing. This, however, is incorrect. For if it were correct, then someone who said "an existing tree," would be combining the very same presentations, and would thus also be expressing a belief in the tree. Yet this is not the case." (Brentano 1928, 42)

8. The existence of objectual attitudes has sometimes been called into question but is ably defended by Forbes (2000) and Montague (2007) among others.

9. There are uses of 'belief in' that may denote a noncognitive attitude, as in 'believe in yourself!' or 'we believe in the future' (which seem to denote emotional attitudes such as confidence and hope). But there is also the cognitive usage highlighted in the main text.

Among attitudes that do not incorporate commitment to *X*'s existence, some expressly involve the opposite commitment, namely to *X*'s *nonexistence*; others are 'existentially silent.' I would love to have a gold-coated private jet; my desire for such a jet, and my contemplation of it, commit me neither to the jet's existence nor to its nonexistence. They are non-committal on the question of the gold-coated jet's existence. By contrast, my belief that Shrek does not exist is not neutral in this way. It *takes a stand* on Shrek's existence—a negative stand. From a Brentanian perspective, this means that there must also be a cognitive objectual attitude that incorporates mental commitment to nonexistence. We may denote this attitude with the expression 'disbelief in': I disbelieve in Shrek in the same sense in which I believe in the Eiffel Tower, and I disbelieve in dragons in the same sense I believe in ducks. What characterizes disbelief-in is that it exhibits the attitudinal property of representing-as-nonexistent.

Brentano's terminology is different. He calls the cognitive objectual attitude that embodies mental commitment to something's existence 'acceptance' or 'acknowledgement' (*Anerkennung*) and the cognitive objectual attitude embodying commitment to nonexistence 'rejection' or 'dismissal' (*Verwerfung*). What matters for our present purposes is not the terminology, but the direction of explanation between existential belief-that and belief-in. Because of a long philosophical tradition of treating propositional attitudes as fundamental in cognition, it is natural for us to analyze 'S believes in X' in terms of 'S believes that X exists.' From a Brentanian perspective, however, this is philosophically misleading. The more fundamental notion is belief-in, precisely because it captures correctly the locus of existence-commitment (as pertaining to the attitude, not content). Accordingly, Brentano would propose to take 'S believes in X' as fundamental and paraphrase 'S believes that X exists' into it (for more on this see Kriegel forthcoming).<sup>10</sup>

Brentano's attitudinal account of mental existence-commitment does raise a problem. If mental existence-commitment is an aspect of existence-committing acts' *content*, then linguistic existence-commitment can be construed in terms of linguistic acts with the very same content. But this cannot work if mental existence-commitment is an aspect of mental acts' *attitude*. A structurally similar account of linguistic existence-commitment would still be possible if there was an existence-committing *force* in lan-

---

10. Szabó (2003) and Textor (2007) also reject the analysis of 'S believes in X' in terms of 'S believes that X exists,' but on different grounds.

guage to parallel the existence-committing *attitude* in thought. But no such force appears to exist. In other words, there is no such thing as *linguistic* representing-as-existent. Perhaps assertion can be thought of as linguistic representing-ad-true, but that is not yet representing-as-existent. So what might existence *talk* amount to if Brentano is right about ‘existence *thought*’?

##### 5. *Linguistic existence-commitment: Brentano’s fitting attitude account*

For Brentano, in asserting ‘X exists,’ we are not saying that X has the property of existing, nor that some X-distinctive properties are instantiated. What we are saying this: that X is a suitable object of acceptance, an appropriate intentional object of belief-in. We are saying that acceptance would be the correct attitude to take toward X, that is, that the right attitude to take toward X is that of believing in it. If X is to be an intentional object of belief-in or disbelief-in, it ought to be the object of belief-in. By contrast, when we say that Y does *not* exist, what we are saying is that if Y is to be an intentional object of belief-in or disbelief-in, it ought to be the object of *disbelief-in*. The correct attitude to take toward Y is that of disbelieving in it. In that sense, Y is a suitable (intentional) object of rejection or disbelief-in.<sup>11</sup>

This account of existence talk can be summarized, or sloganized, with what I will call *Brentano’s Dictum*:

(B1) To be is to be a fitting object of acceptance/belief-in.

Although I formulate Brentano’s dictum in the material mode of speech, it is intended in the first instance as an account of existence *talk*, that is, of linguistic existence-commitment. Note well: in B1, ‘object’ means *intentional object*, not *entity* or *concrete particular*. X is an object of my acceptance in the same sense that my wife is the object of my affection—she is that at which my affection is intentionally directed, and an existent is that at which fitting/correct acceptance is intentionally directed.

---

11. The term Brentano prefers in this context is *richtig*, most naturally translated as ‘correct’ or ‘fitting.’ But in one place he offers a number of synonyms—*konvenient*, *passend*, and *entsprechend* (Brentano 1889, 74)—which are more or less interchangeably translatable as ‘appropriate,’ ‘suitable,’ ‘fitting,’ and ‘adequate.’

Importantly, B1 is not intended as a substantive account of existence. By this I mean that it does not attempt to capture the intrinsic nature of a property of existence. The view is *not* that existence is the property whose nature is being-fittingly-acceptable. In fact, for Brentano *there is no such property as existence*, though there are of course existents.<sup>12</sup> This is precisely why existence-commitment cannot be part of the content of a mental state. There is not some *aspect* of the world, or of things in it, that we may call existence. There is just a *modification* of our awareness of things—a *way* we have of relating to the world (and the things in it) in thought—that we may call existence-commitment. Thus existence talk is in the first instance just a way of describing our awareness or representation of the world, not a way of describing the world itself.

One way to bring this point out is to formulate Brentano's Dictum in overtly contrastive terms. Compare:

(B1a) To be is to be a fitting *rather than unfitting* object of acceptance.

(B1b) To be is to be a fitting object of acceptance *rather than rejection*.

To be sure, B1a is true, insofar as all existents are fitting rather than unfitting objects of acceptance or belief-in. But B1a does not *account* for existence talk, it does not *explain* the function of such talk. What explains that is B1b, the thought that to say that X exists is to take a stand on which attitude it would be correct to take toward X, which attitude is *appropriate* for X.

To that extent, Brentano's account of existence talk can be thought of as a sort of *fitting-attitude account*. Such accounts have recently proliferated in metaethics (see Jacobson 2011). The basic idea is that for X to be good is for it to be a fitting object of approval or the like pro attitude; for X to be bad is for it to be a fitting object of disapproval or the like con attitude. Interestingly, Brentano is commonly admitted to be the first fitting-attitude theorist *of value* (Brentano 1889).<sup>13</sup> Importantly, however, Brentano never

---

12. For example, he writes: 'In calling an object good we are not giving it a material (*sachliches*) predicate, as we do when we call something red or round or warm or thinking. In this respect the expressions good and bad are like the expressions existent and nonexistent. In using the latter, we do not intend to add yet another determining characteristic of the thing in question; we wish rather to say that whoever acknowledges [accepts] a certain thing and rejects another makes a true judgment. And when we call certain objects good and others bad we are merely saying that whoever loves [has a pro attitude to] the former and hates [has a con attitude to] the latter has taken the right stand.' (Brentano 1952, 90)

13. Brentano writes: 'We call a thing *good* when the love relating to it is correct. In the

intended this as a substantive account of the nature of value, but rather as an account of value talk. For him, there is no such property as goodness, though there are goods, just as there is no such property as existence, though there are existents.<sup>14</sup> Thus his accounts of the real and the good are intended to be structurally symmetrical, something he is quite explicit on in several places.<sup>15</sup> Accordingly, we would be quite justified to call Brentano's approach a *fitting-attitude account of linguistic existence-commitment*.

It might be suggested that Brentano's account is a form of 'metaontological expressivism,' since it casts linguistic existence-commitment as a matter of expressing an attitude rather than describing a state of affairs. In *one* sense, this may well be right, since asserting that X exists may be taken to just express the attitude of believing in X. However, this is very different from expressivism as standardly thought of (in metaethics and elsewhere), since the attitude expressed, believing-in, is cognitive rather than conative or emotive. Accordingly, the so-called Frege-Geach Problem does not arise (Geach 1960).<sup>16</sup> Suppose a subject judges both that there is a party and that if there is a party then there is booze, which leads her to judge that there is booze. The validity of this reasoning is captured in traditional *modus ponens*. In Brentano's hands, the reasoning must be recast as follows: the subject both believes in a party and disbelieves in a boozeless party, which leads her to believe in booze. It is true that the traditional formalization of *modus ponens* cannot be used to explain the

---

broadest sense of the term, the good is that which is worthy of love, that which can be loved with a love that is correct.' (Brentano 1889, 18). Likewise: '... everything that can be thought about belongs in one of two classes—either the class of things for which love [pro attitude] is appropriate, or the class of things for which hate [con attitude] is appropriate. Whatever falls into the first class we call good, and whatever falls into the second we call bad.' (Brentano 1930, 21f.)

14. Brentano is clearest on this eliminativist take on goodness in a 1909 letter to Kraus: 'What you seek to gain here with your belief in the existence of goodness with which [pro attitudes] are found to correspond is incomprehensible to me' (Brentano 1966, 207, quoted in Pasquerella 1993, 238; see also Chisholm 1986, 51f.).

15. For discussion of the structural similarity between Brentano's accounts of existence and goodness, see Seron (2008) and Kriegel (forthcoming).

16. The Frege-Geach Problem is often raised for metaethical expressivism. The Problem that if moral statements are not descriptive, they could not play the kind of inferential role that they appear to. For example, one can reason as follows: if my sister did not visit my father in hospital yesterday, then I ought to do so today; she did not; therefore, I ought to. This type of inferential interaction appears to require a descriptive content on ought statements. If we construe moral statements as non-descriptive, and instead as expressive of commendations (say), they would not integrate as well into this kind of inference. For the following seems ill-formed: if my sister did not visit my father in hospital yesterday, then hurrah to visiting him today!; she did not; therefore, hurrah to visiting him today!

validity of this reasoning, but since the reasoning involves only cognitive states, all is needed is a reworking of inference rules within the Brentanian framework. Brentano sketches how to do this in Brentano 1956, but the framework is fully developed only by Peter Simons (1987), who replaces

$$\frac{p \rightarrow q \quad p}{q}$$

with

$$\frac{Na \bar{b} \quad Ea}{Eb}$$

This reads: *a* is not without *b*; *a* is; therefore, *b* is. With this rule in place, we can readily explain why the subject's reasoning to the conclusion that there is booze is valid. Hurrah: the Frege-Geach Problem is avoided!

The key notion in Brentano's fitting attitude account of existence talk is clearly that of acceptance. I now turn to fleshing out this notion.

## 6. *Acceptance*

Arguably, the best way to appreciate Brentano's notion of acceptance is to identify its theoretical role in the architecture of the mind (as Brentano conceives of it). That architecture is in turn best understood in the context of Brentano's lengthy discussion of the classification of mental phenomena. Brentano's terminology here is somewhat dated, so I will briefly present the classification in his own terms, then move to exposition in more modern terminology.

Brentano divides mental phenomena into three 'fundamental classes': presentation (*Vorstellung*), judgment (*Urteil*), and interest (*Interesse*). He also claims that judgment and interest are 'grounded in' presentations, but appears to allow a presentation to also occur without grounding other states. Unlike presentations, judgments and interests are each divided into two opposing kinds: judgment into acceptance (*Anerkennung*) and rejection (*Verwerfung*), interest into love (*Liebe*) and hate (*Hass*).<sup>17</sup> The classification, in these terms, is as follows:

---

17. See Brentano 1874, II Chap. 6–8, including: 'my three classes are not the same as those which are usually proposed. In the absence of more appropriate expressions we designate the

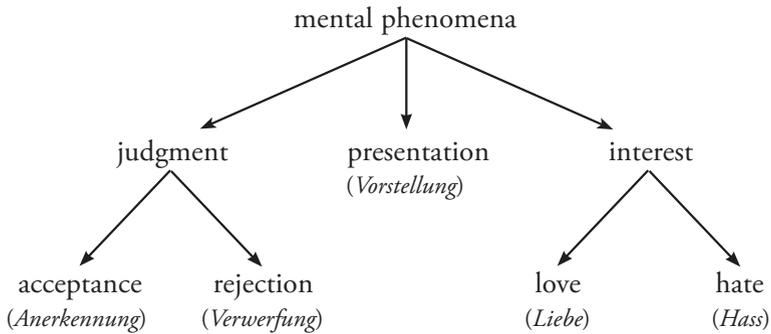


Figure 1. *Brentano's classification, in his terminology*

Some clarifications are called for.

First, Brentano uses the term 'judgment' in an extremely wide sense that covers all mental states with mind-to-world direction of fit (as we would say today).<sup>18</sup> More precisely, a judgment is any state that carries mental commitment to the truth or falsity of what it presents.<sup>19</sup> This includes believing, remembering, speculating, and the like ostensibly propositional attitudes. In addition, however, it includes perceptual states.<sup>20</sup> A visual experience of a yellow lemon has veridicality conditions in the same sense belief has truth conditions. Both are in the business of getting things right. Both have a mind-to-world direction of fit. Accordingly, both are judgments in Brentano's sense.

Secondly, it is significant that for Brentano judgments divide into two categorically different kinds, positive and negative. On this view, to

---

first by the term "presentation," the second by the term "judgment," and the third by the terms "emotion," "interest," or "love." (Brentano 1874, 198) A clearer presentation of the view is this: 'It is clear that all modes of relation to an object fall into three classes: presentation, judgment, and emotion [interest]. The second and third modes always presuppose the first, and in both we find a contrast, in that a judgment is either a belief or a denial, and an emotion is either a form of love or hate.' (Brentano 1928, 42)

18. The distinction between mind-to-world (or 'thetic') direction of fit and world-to-mind (or 'telic') direction of fit is due to Anscombe (1957) and is developed by Searle (1983). The idea is that 'cognitive' mental states, such as belief, are supposed to fit themselves to the way the world is (we want our beliefs to fit the way the world is), whereas 'conative' mental states, such as desire, are supposed to have the world fit them (we want the world to fit the way our desires are).

19. Thus he writes: 'By "judgment" we mean, in accordance with common philosophical usage, acceptance (as true) or rejection (as false)' (Brentano 1874, 198).

20. See, e.g., Brentano 1874, 209.

disbelieve that  $p$  is not just to believe that  $\neg p$ . Nor is it to fail to believe that  $p$ . Rather, it is a *sui generis* attitude irreducible to belief's presence or absence. Even if we accept the law of excluded middle—which not everybody does—there is on this view a psychologically real difference between believing that  $\neg p$  and disbelieving that  $p$ . The former involves mental commitment to the truth of  $\neg p$ , the latter mental commitment to the falsity of  $p$ . This is a psychologically real categorical difference between two kinds of act. In this respect, disbelief parallels displeasure: being displeased that  $p$  is nothing like being pleased that  $\neg p$ .<sup>21</sup> Thus believing and disbelieving can have the same content, and take the judgment attitude toward it, but one takes the positive-judgment attitude while the other takes the negative-judgment attitude. (What motivates this categorical distinction to Brentano is mostly (i) the testimony of his introspection, or rather his 'inner perception,'<sup>22</sup> and (ii) the straightforward account of negative existentials they allow for.<sup>23</sup>)

Thirdly, Brentano's interest category covers a large group of phenomena, including emotion, affect, the will, and pleasure/pain. Desiring that  $p$ , wishing that  $p$ , (dis)approving of  $p$ , being sad that  $p$ , being pained by  $p$ —all these belong to a single fundamental class, according to Brentano.<sup>24</sup> What unifies this class is this: just as states of judgment involve mental commitment to the truth or falsity of what they present, so states of interest involve mental commitment to the goodness or badness of what *they* present.<sup>25,26</sup> We might say that they are all states with world-to-mind direction of fit.

21. It might be objected that in the case of pleasure and displeasure, there is a *phenomenal* difference between the two kinds of acts. But for Brentano, there is also a phenomenal difference between belief and disbelief, at least in the sense that they appear differently to inner perception (which is the only sense of 'phenomenal difference' Brentano would accept).

22. Brentano distinguishes between inner *perception* (*Wahrnehmung*) and inner *observation* (*Beobachtung*), and identifies introspection with the latter (see, e.g., Brentano 1874, 29f.). The distinction is extremely important for Brentano's views on the structure of consciousness and on psychological methodology, but not very important for our present purposes.

23. In addition, Brentano harbors an evident penchant for pervasive symmetries, and the parallelism between positive and negative judgments on the one hand and positive and negative affective states on the other is arguably operative in making the view appealing to him.

24. This is argued for especially in Brentano 1874 II Chap. 8.

25. See, e.g., Brentano's lengthy note on the concepts of truth and existence in Brentano 1889: 73–5. The analogy is particularly explicit in some as yet unpublished lecture notes: 'The good relates to the third class of mental states [interest] as the true to the second [judgment]. Loving is analogous to affirming, hating to denying, rejecting.' (Ms 107c 231, quoted in Seron 2008, 37; see also Ms 107c 236, quoted in Seron 2008, 48f.).

26. Here, the terms 'goodness' and 'badness' are used in the most generic sense. Thus, the relevant notion of the good covers a good car, a good person, a good choice, and so on.

Fourthly, as in the case of judgment, interest divides into a positive kind and a negative kind. ‘Love’ is Brentano’s idiosyncratic term for all states that involve mental commitment to the goodness of what they represent—what we would call today *pro attitudes*. ‘Hate’ is his term for all *con* attitudes. Furthermore, these goodness- and badness-commitments are attitudinal properties. To like cats is not to represent cats as good, but to represent-as-good cats; to dislike raccoons is to represent-as-bad raccoons. Thus we obtain a pleasing symmetry between the cognitive domain and the conative domain, between judgment and interest.

Fifthly, it is not immediately clear that in claiming that judgment and interest are grounded in presentation, or have presentation as their grounds/foundations (*Grundlage*), Brentano has in mind the notion of grounding currently widely discussed.<sup>27</sup> One similarity, however, is that Brentanian grounding is a matter of *ontological asymmetric dependence*, or what Brentano calls ‘unilateral separability’: a mental state can be a presentation without being a judgment, but it cannot be a judgment without being a presentation.<sup>28</sup> (Compare: an animal cannot be a cat without being a mammal but it can be a mammal without being a cat; in that sense cat-ness is ontologically dependent upon mammal-hood.) More specifically, judgment and interest are obtained from presentation by two different attitudinal *modifications* (one embodying mental commitment to truth/falsity, the other embodying commitment to goodness/badness).

Sixthly, presentation can also remain unmodified—modified by neither type of mental commitment. That is how we obtain such states as merely entertaining that *p* or contemplating *p*, as well as phantasms and (day) dreaming. Such states are typically non-committal either on the truth or on the goodness of what they present. They involve neither a mind-to-world direction of fit nor a world-to-mind direction of fit; they are direction-of-fit-less, if you will. Where judgments represent-as-true/false and interests represent-as-good/bad, presentations merely represent. They thus constitute a third species of the genus of which judgment and interest are the other two species. Note that this creates a certain ambiguity in Brentano’s use of the term ‘presentation’: it is used both to denote the genus and to denote one of its species. To remove this ambiguity, we may refer to the species as ‘*mere presentation*.’

27. See Fine (2001) for seminal work on the modern notion.

28. On the notion of unilateral (or ‘one-sided’) separability, see mostly Brentano 1982 Chap. 2.

To summarize, Brentano’s classification divides mental states into three ‘fundamental classes,’ but also into five kinds, insofar as it distinguishes two categorically different kinds of judgment and of interest. Crucially, the difference between all five kinds of state is an *attitudinal* difference—the content can be exactly the same.<sup>29</sup> Thus we get from Brentano five different kinds of *attitude*.<sup>30</sup> A mental state with the content that *p* may either (i) represent-as-true *p*, (ii) represent-as-false *p*, (iii) represent-as-good *p*, (iv) represent-as-bad *p*, or (v) merely-represent *p*. Brentano labels these five kinds of state acceptance, rejection, love, hate, and (mere) presentation. In more modern terminology, we would say that they are (i) positive states with mind-to-world direction of fit (paradigmatically: belief), (ii) negative states with mind-to-world direction of fit (paradigmatically: disbelief), (iii) positive states with world-to-mind direction of fit (paradigmatically: approval), (iv) negative states with world-to-mind direction of fit (paradigmatically: disapproval), and (v) states with no direction of fit (paradigmatically: contemplation). We may thus represent Brentano’s classification, terminologically updated, as follows:

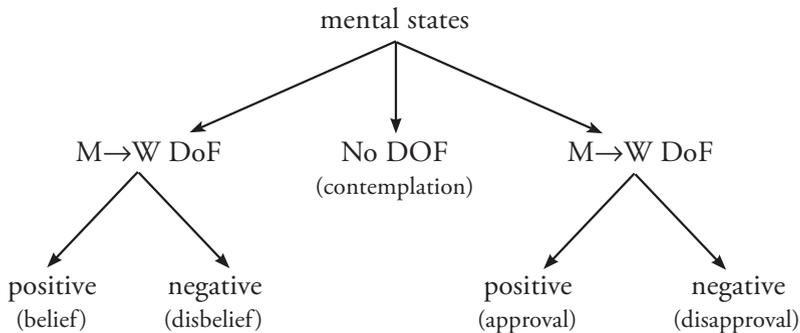


Figure 2. *Brentano’s classification, in modern terminology*

29. Brentano is explicit on the fact that what separates judgment from mere representation is their attitude, or as he puts it their ‘mode of awareness’ (*Weisen des Bewusstseins*). He writes: ‘When we say that presentation and judgment are different fundamental classes of mental phenomena... we mean that that they are two completely different modes of awareness of an object’ (Brentano 1874, 201; my translation).

30. They are five *at the relevant level of abstraction*. At other levels of abstraction they may be even more. For example, Brentano distinguishes between positive judgments that involve commitment to the necessity of what is represented and those that involve commitment to the contingency of what is represented. This too is an attitudinal difference for Brentano (see, e.g., Brentano 1930, 121), but it pertains to a *subspecies* of positive judgment (as well as of negative judgment).

This gives us a first handle on Brentano's notion of acceptance: acceptance is that state which fills the (i) spot in this classification, that is, a positive state with world-to-mind direction-of-fit.

This explains such straightforward statements of Brentano's approach to existence talk as this one (from an 1889 lecture):

Let us say that the area to which affirmative judgment is appropriate is the area of the existent ... and that the area to which the negative judgment is appropriate is the nonexistent. (Brentano 1930, 21)

Accordingly, Brentano's Dictum can be sharpened into the following thesis:

(B2) To be is to be a fitting object of a positive mental state with world-to-mind direction of fit.

This raises an immediate problem, however. In §4, we characterized acceptance as an *objectual* attitude that exhibits the attitudinal property of representing-as-existent. Here we appear to characterize it much more widely, as covering all positive attitudes with a mind-to-world direction of fit. This would appear to include believing-that, remembering-that, and other *propositional* attitudes exhibiting the attitudinal property of representing-as-true. The two characterizations appear to misalign. So which is the correct characterization of acceptance according to Brentano? The answer is that for Brentano this tension or misalignment is merely apparent. For on his view, *all judgments are existential*. Rather than defend this surprising but eminently defensible claim of Brentano's (Kriegel forthcoming), here I wish to show that the fitting attitude account of existence talk does not rely on it.<sup>31</sup>

## 7. *Believing in things*

Regardless of whether all judgments are objectual, the notion of believing in something, as a cognitive objectual attitude embodying existence-commitment, is clearly legitimate. There is clearly a psychologically real mental state corresponding to this description. Isolating it within the classification of mental states should therefore be possible, even if with greater complexity.

---

31. Brentano develops his non-propositional account of judgment most fully in Brentano 1874 II Chap. 7, §§4–8 and §§13–15 (but see also Brentano 1956). The account is explained and motivated in Chisholm (1976) and Kriegel (forthcoming).

Thus, within the genus Brentano calls judgment, and which we identified as mental states with mind-to-world direction of fit, we may distinguish predicative and existential species, with the latter dividing into propositional and objectual subspecies. Predicative beliefs are such as that cats are cute; existential beliefs divide into propositional ones (e.g., the belief that there are cats) and objectual ones (e.g., the belief in cats). If we now focus on existential belief-that and belief-in, we see that each divides into a positive and a negative variety according to the operative attitudinal property. This produces a matrix of four types of existential belief: e.g., believing that there are cats, disbelieving that there are cats, believing in cats, and disbelieving in cats. Thus:

	<u>Positive</u>	<u>Negative</u>
<u>Propositional</u>	represent-as-true (belief-that)	represent-as-false (disbelief-that)
<u>Objectual</u>	represent-as-existent (belief-in)	represent-as-nonexistent (disbelief-in)

Figure 3. *Four types of existential state*

The result is a characterization of belief-in as a positive objectual existential state with mind-to-world direction of fit (see Figure 4 below). We can use this specification to further sharpen Brentano’s Dictum:

(B3) To be is to be a fitting object of a positive objectual existential state with world-to-mind direction of fit.<sup>32</sup>

---

32. Recall that the relevant contrast for Brentano’s Dictum’s explanatory purposes is with other types of attitude. So we could even say: (B3b) To be is to be a fitting object of a positive cognitive objectual mental state with world-to-mind direction of fit *as opposed to other mental states*. The ‘other’ mental states would be negative ones (such as disbelief), non-cognitive ones (such as perception), propositional ones (such as belief-that), or ones with a world-to-mind direction of fit (such as approval).

Mental existence-commitment, on this view, is a matter of taking the attitude encircled in Figure 4 toward something. And linguistic existence-commitment is a matter of asserting the suitability of some intentional object for this kind of mental existence-commitment. Thus, when we say that Obama exists, what we are saying is that toward Obama it is correct to take the attitude encircled in Figure 4. When we say that Shrek does not exist, meanwhile, what we are saying is that Shrek is a fitting intentional object for a *negative* objectual existential state with world-to-mind direction of it.

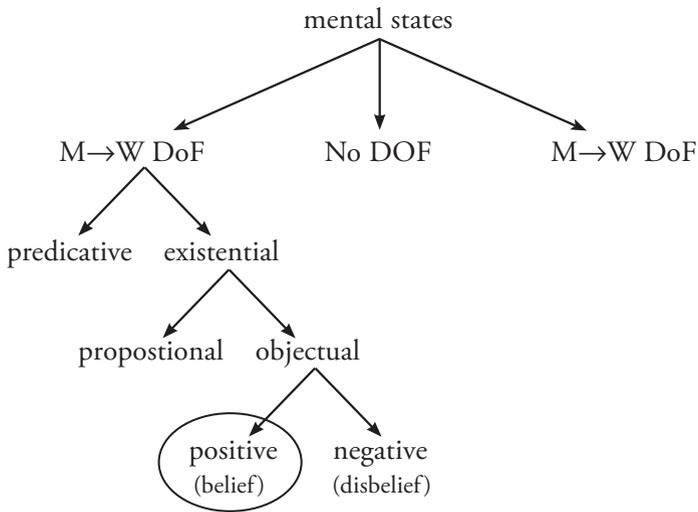


Figure 4. *A quasi-Brentanian classification*

It might be objected that perception is also a positive existential mental state with mind-to-world direction of fit, but is not a form of belief-in. A visual experience of Obama commits to Obama's existence, but does not constitute believing in Obama. Various responses to this objection are possible, but the simplest is to distinguish two subspecies of existential objectual attitude, sensory and cognitive (perceiving X and believing in X), and say that to be is to be a fitting object of a *cognitive* positive objectual existential state with world-to-mind direction of it.

More generally, it may turn out that the present theoretical-role characterization of belief-in is incomplete, in that some other states satisfy it as well. All this would show, however, is that the theoretical role distinctive

of belief-in needs further specification. Ultimately, it is clear that there is such a thing as believing in things, and that it appears somewhere in the complete and correct classification of mental states. An exact characterization of the theoretical role of belief-in may not be trivial, but it is surely available. As long as that is the case, existence talk can be account for in terms of fittingness for belief-in. What matters for our present purposes is that belief-in is (a) psychologically real, (b) existentially committal, and (c) embodies its existence-commitment as part of its attitude, that is, in virtue of exhibiting the attitudinal property of representing-as-existent. If (a)–(c) hold, then the Brentanian approach to existence talk is viable.

#### 8. *Advantages of the fitting attitude account of existence talk*

Brentano's approach is very different from the three more standard accounts. Most importantly, for Brentano mental existence-commitment does not involve attribution of a property to anything. There is a sense in which linguistic existence-commitment does: when we assert that X exists, we implicitly attribute the property of fittingness to the belief in X, indeed X itself is attributed the property of being a suitable object for belief-in. At the same time, this is very different from the property-attribution involved in more standard account of existence talk, insofar as the properties attributed are not ostensibly ontological ones. In any case, in Brentano the parallelism between mental and linguistic existence-commitment is broken: the two look very different, though the account of the latter is derivative from that of the former.

Interestingly, Brentano's unusual approach the main pitfalls stalking the three better-known approaches discussed in §2. To be sure, there may be other solutions to these problems—the literature on this is enormous. But it is remarkable that the problems do not even *arise* for Brentano's fitting-attitude account.

The first view, existence as a substantive first-order property, raised issues with negative existentials, as well as with Hume's observation that the idea of existence 'adds nothing.' Thus, something like 'dragons have the property of not existing' seems to entail, by simple existential generalization, the incoherent-sounding 'there are things that have the property of not existing.' There may be ways around this, but the problem does not even arise in the Brentanian framework. Something like 'it is appropriate to disbelieve in dragons' does not obviously entail 'there are things such that

it is appropriate to disbelieve in them.’ For ‘appropriate to disbelieve in’ may create an intensional context, or at least a context where existential generalization is not supported.<sup>33</sup> In any case, it is clear that Brentano’s approach respects Hume’s ‘adds nothing’ observation: it is precisely because existence is not a *content* feature that there is no content difference between the ideas of a duck and of an existent duck.

The second view, existence as a second-order property, raised issues with direct-reference accounts of proper names, and more deeply recast ‘Obama exists’ as not about Obama (but about his Obamarific properties or the concept OBAMA). In contrast, there is nothing about the fitting-attitude account that requires one to take any position on how ‘Obama’ refers. When one believes in Obama, it is at Obama himself that one’s mental state is directed, not the properties or concept. For the same reason, it is clear that the account respects the intuition that ‘Obama exists’ is about Obama. It is Obama himself who is said to be a fitting object of belief-in.

The third view, existence as a formal first-order property, raised issues with the acquisition of the concept of existence. But Brentano has a straightforward account of how we acquire the concept EXISTENCE. We do not do so by interacting with existents and nonexistents, say with David Chalmers and with Alyosha Karamazov—that would overstretch the notion of interaction. Nor do we do so by interacting with property instantiations and property non-instantiations; when I see a tomato, I interact with a redness instance, but do not in addition interact with a blueness non-instance. According to Brentano, we acquire EXISTENCE by introspectively (or rather inner-perceptually) interacting with mental states that exhibit different attitudinal properties. By inner-perceiving my belief in David Chalmers and my belief in cats, as well as my disbelief in Alyosha Karamazov and my disbelief in ghosts, I interact with a variety of mental states featuring representing-as-existent and representing-as-nonexistent. I also inner-perceive my contemplation of a gold-coated jet, and the four-headed snake phantasm suddenly appearing to me, thereby interacting with the attitudinal property of mere-representing. Through

---

33. Curiously, ‘it is appropriate to believe in’ does seem to support *salva veritate* substitution. Thus, the following seems valid: ‘it is appropriate to believe in Phosphorus; Phosphorus = Hesperus; therefore, it is appropriate to believe in Hesperus.’ There are other cases where an intentional verb supports existential generalization but not substitution. Consider ‘truly believes’: ‘S truly believes that *a* is F, therefore there is an *x* such that *x* is F’ seems valid, but ‘S truly believes that Phosphorus is cool, Phosphorus = Hesperus; therefore, S truly believes that Hesperus is cool’ seems invalid.

such interactions with states that represent-as-existent, represent-as-non-existent, and merely-represent I acquire the concept EXISTENCE (as well as NONEXISTENCE).<sup>34</sup>

Thus the fitting-attitude account of existence talk appears to avoid the problems facing more standard approaches. To repeat, my intention here is not to argue for the *all-things-considered superiority* of Brentano's fitting-attitude account. That would require a much more comprehensive consideration of the objections to the three other views, and of the multitudinous responses to them in the literature. My goal here has been much more modest: to *motivate* consideration of Brentano's alternative approach, as a fourth theory deserving equal attention.

At bottom, however, what motivates Brentano's fitting-attitude account of linguistic existence-commitment is not just the problems facing other views of linguistic existence-commitment. It is also Brentano's attitudinal approach to *mental* existence-commitment. If mental commitment to X's existence comes not from the mental attribution of some property, but from the attitude one mentally takes toward X, then in asserting 'X exists' we are not linguistically attributing a property either, but are commenting on the attitude it would correct or appropriate to take toward X. There is an intriguing shaping of the ontology by (descriptive) psychology here (concordantly with Brentano's general philosophical orientation): it is by attending to the subtle structures of our mental life that we can make sense of ontological discourse.<sup>35</sup>

---

34. Brentano writes: 'Some philosophers have held that this concept [EXISTENCE] cannot be derived from experience. Therefore we shall have to go over this aspect of it in connection with our study of so-called innate ideas. And when we do, we will find that this concept undoubtedly is derived from experience, but from *inner* experience [i.e., inner perception], and we acquire it with reference to judgment.' (Brentano 1874, 210; italics mine) That is, we acquire the concept by inner perception that refers to (is intentionally directed at) judgments, the mental states which exhibit representing-as-existent and representing-as-nonexistent.

35. This work was supported by the French National Research Agency's ANR-11-0001-02 PSL\* and ANR-10-LABX-0087. For comments on a previous draft, I am grateful to David Chalmers, Jonathan Schaffer, Amie Thomasson, and Alberto Voltolini. I have also benefited from presenting this paper at the University of Rennes 1, the University of Liège, and the Jean Nicod Institute. I would like to thank audiences there, in particular Arnaud Dewalque, Filipe Drapeau-Vieira-Contim, Baptiste Le Bihan, Denis Seron, and Mark Textor.

## REFERENCES

- Anscombe, G.E.M. 1957: *Intention*. Oxford: Basil Blackwell.
- Armstrong, David 2004: *Truth and Truthmakers*. Cambridge: Cambridge University Press.
- Brentano, Franz 1862: *On the Several Senses of Being in Aristotle*. Trans. R. George. Berkeley: University of California Press, 1976.
- 1874: *Psychology from Empirical Standpoint*. Edited by Oskar Kraus. Translated by Antos C. Rancurello, D.B. Terrell, & Linda Lopez McAlister. London: Routledge and Kegan Paul, 1973.
- 1889: *The Origin of Our Knowledge of Right and Wrong*. Translated by Roderick M. Chisholm and E.H. Schneewind. London: Routledge and Kegan Paul, 1969.
- 1911: “Appendix to the Classification of Mental Phenomena”. In: Brentano 1874.
- 1930: *The True and the Evident*. Edited by Oskar Kraus. Translated by Roderick M. Chisholm, Ilse Politzer, and Kurt Fischer. London: Routledge, 1966.
- 1933: *The Theory of Categories*. Edited by Alfred Kastil. Translated by Roderick Chisholm and Norbert Guterman. The Hague: Martinus Nijhoff, 1981.
- 1952: *The Foundation and Construction of Ethics*. Translated by E.H. Schneewind. Routledge & Kegan Paul, 1973.
- 1956: *Die Lehre vom richtigen Urteil*. Bern: Francke.
- 1966: *Die Abkehr vom Nichtrealen*. Bern: Francke Verlag.
- 1982: *Descriptive Psychology*. Edited and translated by Benito Müller. London: Routledge, 1995.
- Chisholm, Roderick M. 1976: “Brentano’s Nonpropositional Theory of Judgment”. *Midwest Studies in Philosophy of Mind* (1)1, 91–95.
- 1986: *Brentano and Intrinsic Value*. Cambridge: Cambridge University Press.
- Cummins, Robert 1989: *Meaning and Mental Representation*. Cambridge, Mass: MIT Press.
- Fine, Kit 2001: “The Question of Realism”. *Philosophers’ Imprint* 1(1), 1–30.
- Forbes, Graeme 2000: “Objectual Attitudes”. *Linguistics and Philosophy* 23(2), 141–183.
- Frege, Gottlob 1884: *The Foundations of Arithmetic*. Trans. John L. Austin. Oxford: Blackwell, 1974.
- Geach, Peter Thomas 1960: “Ascriptivism”. *Philosophical Review* 69(2), 221–225.
- Grice, Herbert Paul 1969: “Utterer’s Meaning and Intention”. *Philosophical Review* 78(2), 147–77.
- Hume, David 1740: *Treatise of Human Nature*. Sioux Falls, SD: Nu Vision, 2007.

- van Inwagen, Peter 2003: "Existence, Ontological Commitment, and Fictional Entities". In: Michael Loux & Dean Zimmerman (eds.), *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press, 131–157.
- Jacobson, Daniel 2011: "Fitting Attitude Theories of Value". *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/spr2011/entries/fitting-attitude-theories/>>.
- Kriegel, Uriah 2017 (Forthcoming): "Brentano's Philosophical Program". In: Uriah Kriegel (ed.), *The Routledge Handbook to Brentano and the Brentano School*. London and New York: Routledge.
- (Forthcoming): "Brentano on Judgment as an Objectual Attitude". In: Alex Gzrankowski & Michelle Montague (eds.), *Non-Propositional Attitudes*.
- Kripke, Saul 1972: "Naming and Necessity". In: Donald Davidson & Gilbert Harman (eds.), *Semantics of Natural Language*. Dordrecht: Reidel, 253–355 and 763–769.
- Lowe, Edward Jonathan 2008: "Two Notions of Being: Entity and Essence". *Philosophy* 83(62), 23–48.
- Meinong, Alexius 1904: "The Theory of Objects". Translated by Isaac Levi, D. B. Terrell, & Roderick Chisholm. In: Roderick Chisholm (ed.), *Realism and the Background of Phenomenology*. Atascadero, CA: Ridgeview, 1981, 76–117.
- Montague, Michelle 2007: "Against Propositionalism". *Noûs* 41(3), 503–518.
- Parsons, Terence 1980: *Nonexistent Objects*. New Haven: Yale University Press.
- Pasquerella, Lynn 1993: "Brentano and Aesthetic Intention". *Brentano Studien* 4, 235–249.
- Quine, Willard V.O. 1948: "On What There Is". *Review of Metaphysics* 2(5), 21–38.
- Russell, Bertrand 1905: "Review of A. Meinong, *Untersuchungen zur Gegenstandstheorie und Psychologie*." *Mind* 14, 530–538.
- Schaffer, Jonathan 2009: "On What Grounds What". In: David Chalmers, David Manley & Ryan Wasserman (eds.), *Metametaphysics*. Oxford and New York: Oxford University Press, 347–383.
- Searle, John 1983: *Intentionality*. Cambridge: Cambridge University Press.
- Seron, Denis 2008: "Sur l'analogie entre théorie et pratique chez Brentano". *Bulletin d'analyse phénoménologique* 4(3), 23–51.
- Simons, Peter 1987: "Brentano's Reform of Logic". *Topoi* 6, 25–38.
- Szabó, Zoltán Gendier 2003: "Believing in Things". *Philosophy and Phenomenological Research* 66(3), 584–611.
- Textor, Mark 2007: "Seeing Something and Believing IN It". In: Mary Margaret McCabe & Mark Textor (eds.), *Perspectives on Perception*. Frankfurt: Ontos, 65–87.

- Thomasson, Amie 2015: *Ontology Made Easy*. Oxford and New York: Oxford University Press.
- Voltolini, Alberto 2012: "All the Existences that There Are". *Disputatio* 4(32), 361–383.
- Williamson, Timothy 2002: "Necessary Existents". In: Anthony O'Hear (ed.), *Logic, Thought, and Language*. Cambridge: Cambridge University Press, 233–251.

## COMPUTER-GENERATED MUSIC, AUTHORSHIP, AND WORK IDENTITY

Maria Elisabeth REICHER  
University of Aachen

### *Summary*

In a paper entitled “Computer Composition and Works of Music: Variation on a Theme of Ingarden” (1988), Peter Simons explores some ontological problems that ensue from the use of certain forms of composition software, where the final outcome (the score) is the product of random processes within the computer. Such a method of composition raises, among others, the following questions: What kind of work (if any) has been created? Is it a work of music in the first place? Who is the composer/author? Is it the software programmer, the user, the one who selects a particular score for public performance, the computer? What is the relationship of distinct products of the same programme? Are they instances of the same work? In this paper, I shall re-examine these questions.

### *1. The problem of computer-generated music*

In his paper “Computer Composition and Works of Music” (1988), Peter Simons raises, among other things, some problems concerning a particular kind of computer-generated musical work, i.e., works that are generated by means of a particular kind of composition software. Here software is programmed in such a way that, once the user determines certain parameters and starts the programme, the computer generates a musical score without any further human intervention. Furthermore, the programme has a random generator built in, such that, most probably, if the programme gets started twice, the results will be more or less conspicuously distinct scores.

Simons, in his paper, does not specify how exactly such composition programmes work. In particular, it is not specified which parameters are already fixed by the programme (and thus are not subject to any decisions made by the user), which ones are determined by the user, and which ones are solely determined by random processes. However, this kind of

indeterminacy is not a shortcoming of Simons' theory, since the ontological problems discussed here arise independently of how exactly such programmes are designed. An ontological theory of computer-generated music should ideally embrace all sorts of computer-generated music, not only music generated by means of a particular brand of software. Therefore, not only shall I leave these matters just as indeterminate as Simons does in his paper, but I even leave room for composition programmes that work all by themselves without any input on the user's part. Thus, I shall assume that there is (or, at least, could be) composition software of various kinds, where the programmes may differ, among other things, in the scope left for decisions on the user's part, and in the scope for random determinations. As a limiting case, one may imagine a programme that leaves no decisions whatsoever to the user, such that using the programme in question just consists in clicking on a start-button.

Suppose a user starts a composition programme, and the result is a score  $S_1$ . *Prima facie*, it seems plausible to say that thereby a musical work  $W_1$  came into being, such that  $S_1$  determines  $W_1$ 's identity conditions.<sup>1</sup> Staying with this assumption, suppose further that a user starts the same composition programme again, and the result is a score  $S_2$ , which determines the identity conditions of a work  $W_2$ , and that  $S_1$  is conspicuously distinct from  $S_2$ . Suppose that two performances take place, one in which the performers use  $S_1$ , and another in which the performers use  $S_2$ , and that even a competent and attentive audience would not identify these as performances of one and the same work (assume the audience has not been informed about the history of the performances, in particular not about their common origin). Intuitively, it seems clear that there are two distinct works,  $W_1$  and  $W_2$ , and that, consequently, the two performances are performances of distinct works. However, this raises the question of who, exactly, is the composer of these works.

If we regard the composition programme (together with the hardware in which it is implemented) simply as a tool for the composer (along with, say, music paper, pencils, and composition textbooks), it seems

---

1. Here and in what follows, I presuppose that a musical work is not identical with a score. This point is granted by most (if not all) ontologists of music, no matter whether they consider musical works to be some sort of abstract objects (see, e.g., Ingarden 1962, Wolterstorff 1980, Levinson 1989) or whether they have nominalist leanings (see Goodman 1976, Rossberg 2012, Uidhir 2012). The most obvious reason for this is that the existence of a score is not a necessary condition for the existence of a musical work. (For analogous reasons, I claim that a musical work is not identical with a performance.)

natural to say that, in the case at hand, the user of the programme is the composer.<sup>2</sup> However, this might strike one as counterintuitive from the outset, since starting a computer programme does not require the sort of creativity and skills normally required for the composition of a musical work. Thus, one might be reluctant to award the software user the title “composer” in the first place. Perhaps this worry gets even more pressing if the same piece of software is used twice: Even if one grants that clicking on a start-button can be an act of musical creation, it seems utterly inadequate to regard somebody who just starts the same programme twice as the author of *two* works—for there is no difference whatsoever in the two acts involved.

One can plausibly assume, however, that a lot of creativity and skill have been invested into the creation of the composition programme itself. Granted this, one might be tempted to award the title “composer” not to the user of the programme, but to the programmer. However, the mere investment of *some sort of* creativity and skill on the part of the programmer is insufficient reason to award the title “composer” to her. A considerable amount of creativity and skill may also be involved in constructing musical instruments. Moreover, the making of musical instruments may also have a considerable impact on musical compositions. Yet it seems clear that making a musical instrument is not the same as composing a musical work. Thus, if the programmer can be credited with being a composer, she must bring to bear *the right kind of* creativity and skill; and it is not obvious that this is the case.<sup>3</sup> Furthermore, even if one granted that the creativity of the programmer was of the right sort, this would not resolve the problem of the repeated usage of the same piece of software. True, the programmer might have invested creativity and skill in her programme; but does that make her the composer of a potentially infinite number of musical works? Intuitively, it seems that the programmer has created *at most* one work, not many.

---

2. It has been suggested to me, in personal communication, to consider the programme as analogous to a musical instrument instead of something analogous to paper, pencils and textbooks. It should be noted that to follow this suggestion would not undermine the point that the user can be considered as composer, since musical instruments can be (and, to my knowledge, frequently are) used as devices in the process of musical composition. However, there is at least one problem with the suggested analogy: A musical instrument is a means to generate *performances* of works, and since a score is not a performance, the analogy breaks down with composition software that yields scores instead of sounds.

3. I shall come back to the comparison of instrument making and programming of composition software below; and I shall argue that there is an essential difference between the two.

Thus, this sort of computer-generated music raises a number of distinct, but interrelated questions: What kind of work has been created? How many works have been created? Who is the author of these works? Are distinct scores generated by the same piece of software instances of the same work or instances of distinct works? In what follows, I shall refer to this cluster of questions briefly as *the problem of computer-generated music*.

In his conclusion of the discussion of the problem of computer-generated music, Simons pleads for something that looks like reductionism with respect to musical works, though he couches his view very cautiously:

My own preference would be to avoid the term “work of music” altogether and distinguish between the composer’s *blueprint* [roughly, a score; MR] and its various possible *specifications* in acoustically recognisable types, each performance of which is an *instantiation* of the type and a *realisation* of the blueprint. Each individual performance will be a concrete event and a *concretisation* of the work. (Simons 1988, 152)

I agree with Simons that computer-generated music raises a number of challenging ontological questions. I do not believe, however, that computer-generated music provides a compelling reason to abandon the concept of a musical work altogether. I agree that the terms “type”, “realisation”, and “concretisation” (and their related concepts) are essential for solving the problem of computer-generated music. However, these terms and concepts and their interrelations need further explication—which is not given in Simons’s paper. Moreover, I believe that Simons’s account overestimates the role of the score at the cost of the composer’s intentions.

For these reasons, I would like to reconsider the case and show that an elaborated ontology of art works (works of music and others) has the resources for a viable account of computer-generated works. As noted above, the problem consists in a number of distinct, though interrelated, questions. Here I rehearse them in a slightly different way:

1. Does the process described above yield *any* musical work at all?<sup>4</sup>
2. If the answer to the first question is “yes”, then *how many* musical works are generated?

---

4. Note that the question is not simply whether a *work* has been generated but whether a *musical* work has been generated. For that some work has been generated is quite obvious: The programmer has created a piece of software, and a piece of software is a kind of work (though perhaps not a work of art).

3. Again, if the answer to the first question is “yes”, then who is the author of the work or works in question?

The first question might sound odd, because it seems so obvious that at least one musical work has been generated. After all, we have scores and performances, and members of the audience of the latter would hardly doubt that they have heard performances of musical works. Yet a negative answer to the first question should not be dismissed too swiftly, for it is not absurd to claim that there might be musical sound events without musical works. The audience surely heard musical sound events; but not every musical sound event must be a performance of a musical work. Imagine that, somewhere in the Mesozoic era, the wind whistles through the woods of horsetail trees, bringing about, by coincidence, a melody that sounds like Mozart’s *Little Night Music*. One could argue that this would have been a musical sound event, but not a performance of a work, since there were no works to be performed back then. Therefore, it is at least *prima facie* a serious option to consider computer-generated music to be a case of music without works.

But suppose the answer to the first question is positive: at least one musical work has been generated. Still, that leaves at least two options. We could say that, in the case described above, we have *two* musical works,  $W_1$  and  $W_2$ . Alternatively, we could say that  $W_1$  and  $W_2$  are not distinct works but rather “instances” of one and the same work (where the notion of an *instance* of a work is still to be explicated).

No matter how the second question is answered, if the first question is answered positively, the third question inevitably arises: Who is the composer? The following answers suggest themselves: the programmer, the computer, or the user.<sup>5</sup>

---

5. Two further options might be brought to the fore. The first is that there is no composer at all; the second is that the composers are the members of the audience. In what follows, none of these shall be considered further. Very briefly, here is why: First, the concept of a work includes the concept of authorship. It is a conceptual truth that a work is something that is authored by somebody (using the term “author” in a wide sense, which includes writers, composers, painters, architects etc.). Therefore, the concept of an authorless work is contradictory. Second, the idea that works are authored by recipients is much discussed in literary theory. In the case at hand, however, the recipients play no essential role. For the question here is not whether members of the audience, when listening to a musical sound event, generate musical works, but rather whether musical works are generated by the use of a certain piece of composition software.

## 2. *An outline of an ontology of art*

In my attempt to answer these questions, I shall rely on a theory about the ontology of works of art, including a theory on the concept of authorship. The theory is (just as Simons's) strongly influenced by Roman Ingarden's work on the ontology of art (see Ingarden 1962, 1972), though it does not stick to this theory in all its details. The theory is not only about works of music, but also works of art in general. For present purposes, however, I shall focus on works of music and literature. Here are the basic principles of this theory:<sup>6</sup>

(OA1) In the realm of music and literature, we have to distinguish *works* from *realisations of works*.<sup>7</sup> Works are *types* (or complexes of types). Types are abstract objects that can be realised (instantiated) in many concrete particulars (i.e., types are *universals*).

(OA2) Realisations of works may be quite distinct entities, e.g., particular sound events, loud readings or recitations, actions, mental processes within an individual's consciousness. Some works (namely those that are complexes of types) require for their full realisation the existence of more than one sort of particular entity (e.g., a particular sound event plus a mental process in an individual's consciousness).

(OA3) Each type determines a class of particulars which are the type's instances, i.e., each work determines the class of its realisations. For instance, a musical work may determine a class of sound events, a literary work may determine a class of mental processes, and so on.

(OA4) Despite their abstractness, works are contingent and temporal objects, i.e., works are brought into being (created) at a particular time in history through contingent acts. The creators of works are called *authors*. Each work has at least one author.

(OA5) Properties of a work determining the class of its realisations I call the work's *internal properties*, in contrast to a work's *external proper-*

---

<sup>6</sup> These and related principles have been developed and argued for in detail in Reicher 1998a, 1998b, 2003.

<sup>7</sup> My use of the term "realisation" is slightly different from Simons'. On my use of the term, it would not make sense to say that a score is realised in a performance. In my terminology, only works (and concretisations) can be realised in performances, but not scores. Scores (as types) can be realised in copies of scores, but not in performances. Neither do I stick to Simons's uses of the terms "instantiation" and "concretisation". As I use the term "instantiation", it is equivalent with "realisation". A concretisation, as I use the term, is always a concretisation of a type, and a concretisation of a type is a type itself. Thus, a performance can never be a concretisation in this sense. (For more on concretisations see OA9 below.)

*ties*. An external property is, e.g., the property of having been created in the year 1824, or the property of having been composed by Schubert. An internal property of a musical work is, e.g., its key or its rhythm, an internal property of a poem is, e.g., its rhyme scheme.

(OA6) Authoring (creating) a work essentially means determining the work's internal properties. The author thereby determines the class of the work's realisations. Thus, the creation of a work consists in particular intentional acts, namely acts of decision-making.

(OA7) As a limiting case of the creation of a work one might consider the mere selection of something pre-given as an object to be presented within a certain context. In the fine arts, this is known as *ready-mades* or *found art*, or, if the objects in question are themselves works of art, as *appropriation art*.<sup>8</sup>

(OA8) Works are *incompletely determined*. That is, an author does not determine all the properties of a work's realisations. Each work has *places of indeterminacy*. Consequently, there may be qualitatively distinct realisations of one and the same work: Correct performances of one and the same musical work may sound quite differently, correct performances of one and the same stage play may look quite differently, correct readings of one and the same poem may lead to distinct correct interpretations. The more places of indeterminacy a work has, the bigger the class of its possible realisations. Many works are highly incomplete, and there are no principled limits on the amount of places of indeterminacy.<sup>9</sup>

(OA9) In some arts, notably in music and theatre, there is something in between the work (as it is determined by its primary author) and its performances. In music, this is called *interpretations*, in theatre *productions* or

---

8. The inventor of ready-made art is Marcel Duchamp. In the second decade of the 20<sup>th</sup> century, Duchamp famously presented mundane commodities like a bottle rack, a urinal, or a snow shovel, in the context of art exhibitions, and claimed to have bestowed upon them the status of art works, just through the act of selection. Today, Duchamp's ready-mades are held to be cornerstones of contemporary art. While Duchamp exclusively used artefacts, others applied the idea to natural objects as well (as in certain forms of driftwood art). In so-called appropriation art, artists copy artworks, claiming that new meaning is generated through the re-contextualisation of the originals.

9. Indeterminacy in music becomes particularly salient in improvised music. *Pure* improvised music (i.e., music without any limits to the freedom of the performer to play whatever she wants) is a rare limiting case. I suggest considering pure improvised music as music without works. Most cases of improvised music, however, are not pure improvisations. That is, in most cases there are limits to the freedom of the performer, for instance, a certain melody or a certain sequence of chords that have to be played. In these cases, a work with many places of indeterminacy sets the limits for the improvising performer's freedom.

*stagings*. I call such entities *concretisations*, since they are less incompletely determined than the original work. Concretisations are still types as one and the same interpretation or staging may be realised in many concrete performances. Conductors and performers of musical works, directors and actors of stage plays partly fill in the places of indeterminacy left by the author of the original work. This mental activity is akin to the activity of the author: it also essentially involves making decisions. Therefore, one might regard concretisations as works in their own right, where conductors, performers, directors, and actors are co-authors. The difference is that conductors, performers, directors, and actors do not create from scratch but (usually) stick to the decisions of the primary author and merely add something to them.

(OA10) Works and realisations have to be distinguished from *notations* on the one hand and from what I call *production artefacts* on the other. A notation is something that provides instructions for the production of realisations. Cases in point are scores and stage play scripts, as well as blueprints in architecture. Notations are always in a *notation system*, i.e., a symbol system that has to be understood by the performer who makes use of it. Production artefacts are, for instance, sound carriers. These are entities that can be used for the production of a realisation in an automated way, i.e., without further intentional acts of the sort described above (indeed without any further intentional action, since the “Play” button might be pressed unintentionally as well). Thus, a DVD on which a musical work is stored is not a realisation of the work, but a play of the DVD might be a realisation.

(OA11) A work’s identity is exclusively determined by its internal properties.

### 3. *A solution to the problem of computer-generated music*

Based on OA1 to OA11, my answer to the problem of computer-generated music is as follows: The questions stated above cannot be answered without further specification of the programme design. It all depends on who makes what decisions at which point of the whole process. The process may comprise the following phases:

- (1) the writing of the programme;
- (2) the determination of certain parameters on the user’s part;

- (3) the starting of the programme;
- (4) the running of the programme (where some features of the resulting score are determined by a random generator);
- (5) the selection of a score for presentation.<sup>10</sup>

The key to a solution of the problem of computer-generated music is the concept of authorship explicated above (see OA4 to OA9). Since authoring a work means determining a work's internal properties, for computer-generated music (as for any other art) all depends on whether there have been intentional acts of decision-making of the relevant sort,<sup>11</sup> and, if so, who made these decisions.

Creativity essentially involves the making of decisions. I am aware that the terms "creative" and "creativity" are used in a number of ways often not clearly distinguished from each other, and the claim that making decisions is essential to creativity perhaps does not do justice to all of them. Some might prefer to highlight the role of imagination, the generation of novel ideas, the associative linking up of things that hitherto have not been linked up, etc. Note, however, that I use the terms "creativity" and "creative" with an emphasis on the *ontological* aspect, i.e., on the aspect that creativity is a process through which works come into existence. In fact, imaginative abilities might play an important role in such processes, but it is doubtful whether they are *necessary*, and they are definitely not *sufficient*. For being an artist (or an inventor, for that matter) it does not suffice to have fanciful ideas in mind. There might be quite a long way to go from having an idea to creating a work; and the latter essentially involves making decisions.

Depending on how a composition programme is designed, it may be that a lot of relevant decisions<sup>12</sup> are made in the course of the whole process, or just a few, or, in a limiting case, none at all. Relevant decisions can be made in most of the phases of the whole process, but not in all of them.

---

10. What I call "presentation of a score" here might consist in its publication with a music publisher or in any other sufficiently recognised art world context, as well as in the production of a performance that is based on it.

11. Decisions of the relevant sort are those through which a work's internal properties are determined. For the rest of the paper, when I simply talk about "decisions" or about "relevant decisions", I always mean this sort of decision, i.e., decisions that determine a work's internal properties.

12. In the present context, relevant decisions are decisions that determine the internal properties of a musical work, e.g., decisions concerning key, rhythm, tempo, melody, timbre, instrumentation, etc.

Obviously, no relevant decisions are involved in the mere starting of the programme and in the course of running it. However, relevant decisions can be involved in writing the programme, in the course of determining certain parameters on the user's part, and, finally, when a score is selected for presentation. Of these three phases the first is the programmer's job, the others are the user's job.<sup>13</sup> Thus, it may be that the programmer makes all the relevant decisions, or the user makes them all, or that some of them are made by the programmer and others by the user.

Probably most often, the relevant decisions will be made in several phases of the whole process, but this is not necessarily so. It is even possible that no relevant decisions are made at all. Consequently, the whole process may generate no musical work at all, or just one musical work, or several; and it may be that there is no composer at all, or that the programmer is the composer, or the user, or programmer and user together. Identifying the computer as a composer is the only option of the ones mentioned above that can be ruled out immediately. Given that a computer does not make decisions but just conducts algorithms determined by the programmer and/or user, the computer cannot be the author of anything.

Thus far I have provided an outline of my answers to the problem of computer-generated music. In what follows, I shall defend them against several objections, thereby spelling them out in more detail.

### *The programmer as composer*

From at least one point of view, it might seem inadequate to credit the programmer with being a composer. After all, the programmer's work is done before a musical work comes into existence, or so it seems. However, writing a composition programme may (and, probably, usually does) involve making decisions of the kind involved in composing a musical work in a more traditional way, e.g., decisions concerning key, rhythm, acoustic colour, instrumentation, etc. The programmer may determine these and other parameters, or she may determine some of them, and leave others open (to be determined either by the user or by the random generator).

---

13. The person who determines certain parameters before the programme gets started is not necessarily identical with the person who selects a score for presentation. I call both of them "user". Thus, it may be that one user determines certain parameters before the programme gets started and another selects a score for presentation. For the sake of simplicity, I often simply talk about "the user"; but it should be kept in mind that it is possible that more than one user is involved.

Furthermore, if she does not fully determine certain parameters, she may still consciously narrow down the class of possible choices regarding these parameters to any extent whatsoever. For instance, the programmer may restrict the possible choice of keys to major and minor, thereby excluding modal keys (which were used in ancient and medieval music as well as in some contemporary music) as well as atonal scales (used by Arnold Schönberg and Anton Webern, among others). The more conscious decisions of the relevant sort the programmer makes, the closer she comes to composing a musical work in the usual sense.

In what follows, I shall discuss some objections against considering the programmer of a piece of composition software as composer of a musical work.

*Objection 1:* The programmer cannot have created a musical work, because her activity—taken by itself—neither generates a score nor a sound event.

*Reply:* The generation of a score and/or sound event is not a necessary condition for the creation of a musical work. The creation of a musical work is essentially a mental process. The writing down of a score and/or public performance are just means to make the results of this process publicly accessible and to preserve them for the posterity.

*Objection 2:* Composing a musical work in the traditional sense involves having a quite definite idea of what a correct performance of the work would sound like. The composer must be able to imagine realisations of her work (perhaps with the help of expedients like the playing of a melody or a sequence of chords on the piano). But the programmer of a piece of composition software of the kind described above is unable to imagine realisations of her work.

*Reply:* Perhaps it is true that *usually* composers are able to imagine realisations of their works. But this does not imply that having an idea of a work's performance in mind is an *essential* ingredient of composing a musical work. The essential ingredient is the making of decisions that narrow down the class of possible performances. Imaginative processes in the composer's mind (if such exist) are mere accidental attachment.

Furthermore, it is not clear that programmers of composition software *never* have *any* idea of what a performance of their work would sound like. Most likely, they are not able to imagine *every* possible performance of their work, but they might well be able to imagine *some* of them. I suspect the same holds true for the majority of works of classical music (that have been composed in the usual way) as well.

*Objection 3:* Even though distinct performances of one and the same work of traditional music may sound quite differently, it is possible (at least for trained persons) to *recognize* the distinct performances as performances of one and the same work. But this does not hold for computer-generated music anymore. (Cf. Simons 1988, 151)

*Reply:* Our capacity to recognize similarities is irrelevant for ontological questions. What matters is the existence of objective similarities. Furthermore, whether an audience would identify two performances as performances of one and the same work does not depend on whether the work is computer-generated or brought into existence by other means. On the one hand, there may be performances of computer-generated works that are easily recognized as performances of the same work; on the other hand, there may be performances of a hand-written work that are not recognized as performances of the same work.

*Objection 4:* If the programmer is to be considered as a composer, since she narrows down the class of possible performances through relevant decisions, don't we have to credit also instrument makers with being composers? For surely, the instrument maker also narrows down the class of possible performances through his decisions.<sup>14</sup>

At first sight, the analogy between programmer and instrument maker looks plausible, but closer inspection shows that it does not hold. Composing a work is to define the work's essence, i.e., it is determining which qualities any realisation of the work necessarily has. It is in this sense that the composer narrows down the class of possible performances of her work; and—as I have argued above—programmers may do exactly this as well. The instrument maker, however, does not do anything of the sort. The instrument maker does not—by determining particular features of his instrument—define any musical work's essence. Of course, this is not to deny that the design of musical instruments may have a considerable impact on the composition of musical works. Since composers (usually) like to have their works performed, they have to take into account which sound structures can actually be realised, and in what ways. Newly designed instruments widen the class of realisable sound structures and thus widen the class of performable works. Particular features of instruments may guide and inspire composers. Thus, the decisions of an instrument maker may have significant impact on those decisions that constitute the composition of musical works; but still

---

14. This objection was brought to my attention by Georg Friedrich.

the instrument maker's decisions do not constitute the composition of musical works.

*Objection 5:* The claim that the programmer is a composer (i.e., the author of a musical work) is based on confusion. The programmer indeed creates something, namely a composition programme, but a composition programme is not a musical work. Perhaps one might also say that the programmer creates a set of instructions, namely a set of instructions for the generation of distinct scores; but sets of instructions, too, are not musical works.

*Reply:* I agree that a composition programme has to be distinguished from a musical work, although a full-fledged ontology of computer programmes is still lacking.<sup>15</sup> From what I know, among computer scientists, it is common to identify a programme not with algorithms (i.e., sets of non-ambiguous rules for the computer) nor with computational processes (i.e., processes that are fully defined by relations between initial states, inputs, and final states), but rather with a text, written in a particular programming language: the so-called *source code*. As I see it, this opens up at least two attractive options for an ontology of computer programmes: programmes as *source code types* or programmes as *computational process types*. I will not take a definite stance on these matters here, but it is clear that a musical work is neither a source code type nor a computational process type. Yet that the programmer creates a piece of software, and that a piece of software is not a musical work, does not entail that the programmer is not also creating a musical work by making the relevant decisions in the course of creating the composition programme.

I also agree that it makes sense to say that the programmer is not only the creator of the programme, but also the creator of a set of instructions, since usually creators of artefacts not only determine the artefact's internal properties but also suggest a way (or perhaps a number of ways) the artefact should be used. This holds in particular for technical artefacts,<sup>16</sup> and surely a computer programme is just that. Instructions for using a piece of software may be very simple ("Click on the start-button!") or much more complex (depending on how much room is left for decisions on the user's part).

Furthermore, I agree that a set of instructions for the generation of a musical work is not itself a musical work. Rather, a set of instructions

---

15. There has not been much research on this topic yet. For two instructive and interesting beginnings see Lopes 2010, Chapter 4, 63f., and Irmak 2012.

16. For a theory of technical artefacts and their functions see Houkes & Vermaas 2010.

can be considered a work in its own right. *Games*, in particular, can be constructed plausibly as sets of instructions (where in some, but not all, cases the instructions involve the use of particular artefacts, like game boards and tiles, balls, rackets etc.—or computers and computer programmes); and surely, games are authored artefacts. But again, this does not rule out that programmers of composition software in addition create musical works.

*Objection 6:* Even though the programmer makes a number of decisions that are similar to those made by composers, the process of programming a piece of composition software is *not creative enough* to be counted as the creation of a musical work.

*Reply:* Admittedly, the decision to restrict the class of available keys to major and minor is not a terribly impressive example of creativity on its own (even if it is made with full consciousness, being aware of the existence of alternatives); the same would hold for the determination of a particular key, say, C minor. Yet it seems to me that the construction of the most impressive products of human creativity can be split up into the making of such mundane decisions. Therefore, insofar the programmer makes decisions of the relevant kind, one might consider her as composing a musical work. Moreover, creativity comes in degrees; and, from an ontological point of view, it seems *ad hoc* to demand a particular amount of creativity as a necessary condition of work creation.

*Objection 7:* Granting that the programmer brings into existence a *musical type* by making certain decisions, this type is nevertheless much too “thin” to be counted as a *musical work* in the proper sense, that is, it has too many places of indeterminacy. Imagine a very primitive composition programme that is governed solely by the following rule: Generate a score such that its performance is a sequence of tones (of arbitrary length and complexity) out of a range of tones so that each tone is characterized by one of 30 pitches, 5 durations, 10 loudnesses, and 20 timbres! Such a programme is likely to generate tons of very dull scores, e.g.:

$S_1$ : a notation of a sound event that consists of a single tone in pitch<sub>1</sub>, duration<sub>1</sub>, loudness<sub>1</sub>, and timbre<sub>1</sub>;

$S_2$ : a notation of a sound event that is like  $S_1$  except that the single tone is repeated, say, 10 times;

$S_3$ : a sound event that is like  $S_2$  except that after the 10 occurrences of the single tone another tone is repeated 10 times which differs from the first one in its duration.

And so forth.

One might admit that the programmer of this piece of composition software created a musical type. After all, she selected pitches, durations, loudnesses, and timbres, thereby setting limits and narrowing down the class of possible performances. However, this type is very thin, i.e., it is much more incompletely determined than usual examples of musical works. Therefore, one should better refrain from calling it a *work* in the first place.

*Reply:* For practical purposes, it might be reasonable to demand a certain amount of “thickness” of something to be granted the status of a work.<sup>17</sup> But from an ontological standpoint, it is doubtful whether gradual differences in the amount of indeterminacy entail a categorical difference between works and non-works.<sup>18</sup> At any rate, the claim that only types of a certain thickness are works would raise the problem of determining what amount of thickness is required for the status of a work; and there is hardly a chance to answer this question in a clear-cut and non-arbitrary way. Therefore, I prefer to grant work status even to very thin types.

What the author of our primitive composition programme is doing is similar to what a composer is doing who gives instructions to a performer for a *restricted improvisation*: “Play whatever you want, but you have to use the following set of pitches, durations, loudnesses, and timbres: ...—and you must not use anything else.” I see no principled reason to deny that such an instruction is a sort of notation of a musical work.

Furthermore, the fictitious composition programme described above is very primitive. If a composition programme should be likely to generate more interesting scores, either the programmer has to make more decisions,

---

17. This suggestion comes quite close to the basic idea of the concept of *Werkhöhe* (also called *Schöpfungshöhe* and *Gestaltungshöhe*) in German copyright law. The intention is that not every product of human creativity (in the wide sense of creativity that I adopted here) should be protected by copyright law, but only those surmounting a certain threshold of originality and/or individuality. Copyright regulations of other countries incorporate similar concepts. On the one hand, this seems to be a reasonable policy, for excessive copyright protection is likely to impede innovation. On the other hand, the concept of *Werkhöhe* is not only notoriously vague, but also ambiguous, for the concepts of originality and individuality oscillate between ingenuity and “thickness”. In other words, both the novelty of a single idea and the complexity of an arrangement of pre-given elements might give a type the required *Werkhöhe*. As a matter of fact, the courts tend to set the threshold pretty low.

18. In light of the difficulties to set an adequate threshold for *Werkhöhe* (see the former footnote), one might argue that even for practical purposes this approach is not useful. Instead of a strict division into works and non-works, based on the concept of *Werkhöhe*, which results in all-or-nothing decisions, it might be better to develop a more fine-grained system of fair use and compensation regulations that take into account gradual differences.

or the programme has to allow for more decisions to be made by the user. The more decisions are made, the more restrictions are built in, the more the class of possible performances is narrowed down, the thicker is the work. Thus, it is imaginable that the activity of the programmer comes close to composing a musical work in the usual sense and that its results are quite as thick as those of more traditional composition processes.

Finally, I want to stress that nothing claimed in the present paper entails that creating a work requires extraordinary skills or artistic giftedness; the creation of a musical work—as it is understood here—is not necessarily an extraordinary achievement. From an ontological point of view, it is better to free the concepts of authorship, creation, work, etc., from the air of ingenuity adhering to them in everyday usage. As I use the terms “work”, “composer”, “author” (etc.) here, their meaning contains only a very small amount of valuation (if any). To say that something is a work is far from saying that it is a valuable, interesting, or original work. For instance, it does not imply that it is worthwhile in any way to listen to a performance of it. Works—both computer-generated and others—may be boring, derivative, trite. To say that something is a work or that a person is an author is not a value judgement.

### *The user as composer*

As I said above, decisions on the user's part can be involved in two phases of the whole process: The user can make decisions either by determining certain parameters before the programme gets started (if the programme allows for that) or when he selects a score for presentation.

When the user makes decisions by determining parameters before the programme gets started, his decisions are similar to the sort of decisions traditional composers make. The only significant difference is that, most likely, the user has to choose among a limited range of choices. He makes certain choices within the range of choices selected by the programmer. He may select a certain key, a certain beat, perhaps even a certain style (Mozart style, or Bach style, or Steve Reich style, etc.). In this case, he makes decisions of the kind involved in composing a musical work, though within the limits set by the programmer. Thus, in this case one may consider the user as creator of a musical work.

If the programme disallows the determination of any parameters on the user's part before it gets started, then such decisions can become involved only after the computing process is finished. The user may only decide

whether a given score should be discarded or presented. This selection for presentation might be considered as the sort of decision involved in authoring a work.

Some of the same objections against the suggestion to consider the programmer as composer might be raised against the suggestion to consider the user as composer as well. This holds in particular for the “not enough creativity” and the “too thin to be a work” objections. The suitable replies are the same in both cases. To rehearse them very briefly: Creativity comes in degrees. There is no objective thickness-threshold for works. The concepts of work and authorship should be understood in a non-evaluative sense.

Besides, it cannot be claimed *a priori* that a user of a composition programme invests less proficiency and efforts in the creation of a musical work than a composer of the traditional kind. Admittedly, composition programmes may make it easier to compose musical works. But this does not entail any particular value judgements on the achievements of a particular composer, and even less on the value of a particular musical work.

The creation of a musical work by selecting a score for presentation is certainly a limiting case of authoring, just as the authoring of ready-mades in the fine arts (see OA7 in Section 2 above). But if one accepts ready-mades as works (as seems common nowadays in the art world), then there is no principled reason to deny this status to selected computer-generated scores. Analogously, if one accepts Marcel Duchamp as the author of the famous *Bottle Rack*, there is no principled reason to deny a user who selects a computer-generated score for presentation the status of an author.

### *No author, no work*

I have argued that *usually* the programming of composition software involves the making of relevant decisions; but this is not *necessarily* the case. One might imagine a programmer who simply compiles all the musical “material” she can gather (all available keys, all rhythms, all acoustic colours etc.) such that no restriction whatsoever is made on her part. This would be an achievement of its own, but it would not be *creative*. Gathering material with the intention to generate a compilation as complete as possible is not to narrow down the class of possibilities. To the contrary: the attempt is to widen the scope of possibilities as far as one can.

It might be a programmer’s intention to do just that: provide a maximum amount of possibilities, to enable everything whatsoever that could

be counted among the realm of music in the widest possible sense. In that case, there would be no reason to consider the programmer as composer. Consequently, the programmer would not have created a musical work.

Now, imagine a programme whose design not only does not involve any relevant decisions, but which, in addition, does not allow for the determination of any parameters on the user's part before it gets started. In that case, if a score gets generated, its relevant features are in no way determined by any decisions of a conscious subject. It is a mere random product. There is no author, and, consequently, no work. The case is analogous to a "score" that becomes generated through random processes in nature (say, water and wind dispersing seashells, pebbles and sprigs in the sands of a lonely beach). Still, there is the possibility that the random product becomes a work through the intentional act of selection for presentation (as a piece of *found art*). But if even this act does not happen, there is no author and, consequently, no work.

### *Shared authorship*

I have argued that the programmer should be granted the status of authorship if the programming involves relevant decisions. I have argued further that the user should be granted the status of authorship if he makes relevant decisions, either in determining parameters before starting the programme or in selecting a score for presentation. Consequently, if such decisions are made, either on the programmer's or on the user's part, a musical work has been created.

That the programmer is the sole author of the generated work is possible. This is the case if the programmer has made decisions of the relevant kind in the writing of the programme, and if there is no relevant impact on the user's part whatsoever (neither by determining parameters before the programme gets started nor afterwards, by selection of a score for presentation).

On the other hand, it is possible that the user is the sole author of the generated work. This is the case if the writing of the programme did not involve any relevant decisions, but the user makes relevant decisions, either by determining certain parameters before the programme gets started, or afterwards, by selecting a score for presentation.

However, I take it that these are special cases that happen rarely, if ever. In the usual case, I suppose, neither the programmer nor the user can be credited with *exclusive* authorship. Rather, the resulting work is the

product of *shared authorship* that involves both the programmer and the user. An author who creates a work has a share in all the work's concretisations (no matter who creates the concretisations), because the author of the original work has determined some of the internal properties of all its concretisations.<sup>19</sup>

Note that *shared authorship* does not necessarily involve collective intentionality as commonly understood (i.e., as involving a sort of mutual agreement). The programmer need not have any beliefs about or expectations on the users. Neither do the users need to have any beliefs about the programmer's intentions. There is no sort of contract between programmer and users, no mutual obligations, etc. This distinguishes shared authorship from traditional forms of collaborative work creation.<sup>20</sup>

### *Concretisations and identity*

I have argued that the programmer may be composer of a musical work (given that she makes relevant decisions in designing the programme). Furthermore, I have argued that the user may be composer of a musical work, if he makes relevant decisions, either by determining certain parameters before the programme gets started, or by selecting a score for presentation. Finally, I have argued that musical works may be co-productions of programmers and users.

In what remains, I shall refer to the work that has been created merely through the creation of the composition programme as  $W_0$ . It has been argued that  $W_0$  indeed is a musical work, and that its composer is the programmer. Now let's get back to the original assumption that the same programme is run twice, and that the first running results in a score  $S_1$ , the second in a score  $S_2$ , such that  $S_1$  and  $S_2$  are substantially distinct. Here I shall tackle the question of the relationship between  $W_0$ ,  $S_1$  and  $S_2$ . Are there works  $W_1$  and  $W_2$  corresponding to  $S_1$  and  $S_2$ ? If so, are they distinct works or "instances" of one and the same work (perhaps "instances" of  $W_0$ )?—My answers are as follows:

First of all, I want to stress that a mere automated generation of a score cannot be the creation of a new work, not even of a concretisation. This

---

19. For an explication of the notion of a concretisation see OA9 in Section 2 above; for an application see the following sub-section.

20. To highlight this difference (which was brought to my attention by Raoul Bussmann), I have chosen the term "shared authorship" instead of "collective authorship" or "collaborative authorship".

is a consequence of OA6, which says that authoring a work essentially means determining a work's internal properties by making decisions that narrow down the class of possible performances. Thereby, it does not matter whether the programme is started by the computer itself (without any further input), or whether the programme is started unintentionally (for instance, by a cat playing around), or whether a user starts it intentionally. In none of these cases is anybody making decisions that determine a work's internal properties (or, more specifically, fill in spaces of indeterminacy in  $W_0$ ). Therefore, if the programmer makes all of the relevant decisions, there is only one work ( $W_0$ ), no matter how many distinct scores are generated by the programme.

But suppose a user determines a set of parameters (in addition to some that have already been determined by the programmer) before the programme gets started. In that case, the user creates a work  $W_1$  such that  $W_1$  is a *concretisation* of  $W_0$ , that is,  $W_1$  has all the internal properties of  $W_0$ , and some more.<sup>21</sup> In other words,  $W_1$  is just like  $W_0$ , except that in  $W_1$  some of  $W_0$ 's places of indeterminacy are filled in, or at least scaled down.<sup>22</sup> Both  $W_0$  and  $W_1$  are incompletely determined, but  $W_1$  is less incompletely determined than  $W_0$ .  $W_1$  is jointly authored both by the programmer and by the user, because  $W_1$  is based upon  $W_0$ , which has been authored by the programmer.

Suppose that another user of the same programme determines a distinct set of parameters. In doing this, he creates a work  $W_2$  which is distinct from  $W_1$  but is, just as  $W_1$ , a concretisation of  $W_0$ . Again,  $W_2$  is less incompletely determined than  $W_0$ ; and again,  $W_2$  is jointly authored by the programmer and by the user. Thus, in this case,  $W_1$  and  $W_2$  are distinct concretisations of the same work  $W_0$ .

Of course, the programme may generate a number of distinct scores based on  $W_1$ : The user may save his selection of parameters and start the programme with this selection of parameters again and again. But, as has been stated already, the mere automated generation of scores does not generate new works.

---

21. Note that the user who creates  $W_1$  is not necessarily the person who started the programme and thus caused the automatic generation of the score.

22. That places of indeterminacy are not filled in but only scaled down means that the number of possibilities is reduced, but that there is still more than one possibility left. Imagine, for instance, that the programme allows for all major and minor keys and that the user rules out the major keys before he starts the programme. Thereby, the user has scaled down a place of indeterminacy without filling it in. For the latter, it would be necessary to determine a *particular* key, say, C minor.

Matters become importantly different, however, if the user takes a selection of the automatically generated scores. Suppose (i) the user determines a set of parameters, thereby creating a work  $W_1$  (which is a concretisation of  $W_0$ ), (ii) the programme generates, say, 100 distinct scores based on this set of parameters, and (iii) the user selects two of these for presentation. I suggest that through these acts of selection two further works are created, say,  $W_{1a}$  and  $W_{1b}$ .  $W_{1a}$  and  $W_{1b}$  are distinct concretisations of  $W_1$ , that is, they are like  $W_1$ , except that some of the places of indeterminacy of  $W_1$  are filled in (or at least scaled down) in  $W_{1a}$  and  $W_{1b}$ . Thus,  $W_1$  is less incompletely determined than  $W_0$  but more incompletely determined than  $W_{1a}$ .

$W_{1a}$  and  $W_{1b}$  are not only concretisations of  $W_1$ , but also concretisations of  $W_0$ ; the relation of *being a concretisation of* is transitive. Concretisations can be “nested”: A work  $W_0$  may have a number of concretisations  $C_1, C_2, \dots, C_n$ , such that:  $C_1$  is a concretisation of  $W_0$ ,  $C_2$  is a concretisation of  $C_1$  and  $W_0$ ,  $C_3$  is a concretisation of  $C_2, C_1$ , and  $W_0$ , and so forth.  $W_{1a}$  (as well as  $W_{1b}$ ) might be used as a starting point for a further concretisation—a musical interpretation—by a conductor or performer. The conductor’s/performer’s concretisation would then be a concretisation of  $W_{1a}$  as well as of  $W_1$  and  $W_0$ .

Now imagine a process like the one just described, except that step (i) is dispensed with—i.e., the user does not determine anything before the programme gets started but selects afterwards two scores out of 100. I still maintain that two distinct concretisations of  $W_0$  are created— $W_{1a}$  and  $W_{1b}$ . In contrast to the case from above, however, here there is no “intermediate” concretisation  $W_1$ .

In a slightly distinct case, it is a user  $U_1$  who selects a score  $S_1$ , and a distinct user  $U_2$  who selects a score  $S_2$ . In this case,  $U_1$  is co-author of  $W_{1a}$ , and  $U_2$  is co-author of  $W_{1b}$ . In case there is the intermediate step (i)—i.e., the determination of parameters before the programme gets started—the final work  $W_{1a}$  may have three co-authors: the programmer, the first user (who determines the parameters before the programme gets started), and the second user (the one who selects the score of  $W_{1a}$ ).

It might seem that the user’s creative impact is bigger when he determines parameters before starting the programme than it is when he just selects particular items from the programme’s output. However, whether this is really the case depends on how the programme is designed. The degree of creative impact (as I use the term here) depends on the extent to which the user’s decisions narrow down the class of possible realisations. Suppose that the programme has a very large random scope, and the user

selects two scores among a huge heap of automatically generated scores. It may be that this user's impact on the class of possible realisations is bigger than the impact of a user who has chosen between a very limited number of options before the programme got started. In either case, however, the user is not the sole author of the resulting works. In both cases, the works are co-authored by programmer and user. When the programmer has created a work  $W_0$ , then she has a share in the creation of all of  $W_0$ 's concretisations. After all, in the cases described above, the programmer has determined *some* of the internal properties of  $W_1$ ,  $W_2$ ,  $W_{1a}$ ,  $W_{1b}$ , etc. However, since she has not determined *all* of the internal properties of these works, she is not the sole author of  $W_1$ ,  $W_2$  etc.  $W_1$  and  $W_2$ , as concretisations of  $W_0$ , can be considered to be co-productions of the programmer and the user—given that there is some impact on the user's part (i.e., some relevant decisions on the user's part). This holds independently of whether the user determines some parameters before the programme gets started or whether he selects among a number of automatically generated scores, or both.

It should be clear by now that  $W_0$  is not identical to any of its concretisations, and that  $W_1$  is not identical to  $W_2$ , etc. For the set of  $W_0$ 's internal properties is distinct from the set of the internal properties of both  $W_1$  and  $W_2$ , and the set of  $W_1$ 's internal properties is distinct from the set of  $W_2$ 's internal properties. Yet the relationship between  $W_0$ ,  $W_1$ , and  $W_2$  is one of *partial identity*.  $W_0$  might be considered a logical part of both  $W_1$  and  $W_2$ , which is to say that the set of  $W_0$ 's internal properties is a subset of the set of  $W_1$ 's internal properties as well as a subset of the set of  $W_2$ 's internal properties.

### *Summary*

The problem of computer-generated music may be couched in the following three questions: Does the process bring about any musical work at all? If yes, how many musical works are brought about, and who is the author of the work or works in question?

The answers I have suggested rest on the idea that to create a musical work essentially consists in determining the work's internal properties, and that this determination is an act of decision-making. Therefore, computers cannot create musical works (given that computers do not make decisions). Whether the process yields any musical work depends on whether any relevant decisions are made. In principle, such decisions can be made by the programmer in designing the programme, and by a user who determines

certain parameters before the programme gets started, and by a user who selects a score for presentation. Given that the programmer makes relevant decisions, she is composer of a musical work (albeit perhaps of the very thin variety). Based on this, users may create works of their own by filling in or scaling down some of the places of indeterminacy in the programmer's work. The resulting works are concretisations of the programmer's work. Such concretisations are products of shared authorship, i.e., they are co-authored by the programmer and the user. The relationship between the programmer's work and its concretisations is one of partial identity.<sup>23</sup>

#### LITERATURE

- Goodman, Nelson 1976: *Languages of Art. An Approach to a Theory of Symbols*. Indianapolis: Hackett.
- Houkes, Wybo & Pieter E. Vermaas 2010: *Technical Functions. On the Use and Design of Artefacts*. Dordrecht et al.: Springer.
- Ingarden, Roman 1962: *Untersuchungen zur Ontologie der Kunst. Musikwerk—Bild—Architektur—Film*. Tübingen: Niemeyer. [English translation: *Ontology of the Work of Art: the Musical Work, the Picture, the Architectural Work, the Film*. Athens: Ohio University Press 1989.]
- Ingarden, Roman 1972: *Das literarische Kunstwerk*. 4<sup>th</sup> edition. Tübingen: Niemeyer. [English translation: *The Literary Work of Art*, Evanston, Illinois: Northwestern University Press, 1973.]
- Irmak, Nurbay 2012: "Software Is an Abstract Artifact". *Grazer Philosophische Studien* 86, 55–71.
- Levinson, Jerrold 1980: "What a Musical Work Is". *Journal of Philosophy* 77 (1), 5–28.
- Lopes, Dominic McIver 2010: *A Philosophy of Computer Art*. Oxon–New York: Routledge.
- Reicher, Maria E. 1998a: *Zur Metaphysik der Kunst. Eine logisch-ontologische Untersuchung des Werkbegriffs*. Graz: dbv-Verlag.
- 1998b: "Works and Realizations". In: Nicola Guarino (ed.), *Formal Ontology in Information Systems. Proceedings of the First International Conference (FOIS'98)*,

---

23 For very helpful comments on a preliminary version of this paper, I wish to thank the participants in the *Forschungskolloquium* of the Philosophical Department at the RWTH Aachen University, Joachim Bromand, Raul Bussmann, Christoph Diehl, Georg Friedrich, Wulf Kellerwessel, Nicola Mößner, and Rochus Sowa.

- June 6-8, Trento, Italy*. *Frontiers in Artificial Intelligence and Applications* 46, Amsterdam et al: IOS-Press, 121–132.
- 2003: “Eine Typenontologie der Kunst”. In: Reinold Schmücker (ed.), *Identität und Existenz. Studien zur Ontologie der Kunst*. Paderborn: Mentis, 180–199.
- Rossberg; Marcus 2012: “Destroying Artworks”. In: Christy Mag Uidhir (ed.), *Art and Abstract Objects*. Oxford: Oxford University Press, 62–83.
- Simons, Peter M. 1988: “Computer Composition and Works of Music: Variation on a Theme of Ingarden”. *Journal of the British Society for Phenomenology* 19, 141–154.
- Uidhir, Christy Mag 2012: “Introduction: Art, Metaphysics, and the Paradox of Standards”. In: Christy Mag Uidhir (ed.), *Art and Abstract Objects*. Oxford: Oxford University Press, 1–26.
- Wolterstorff, Nicholas 1980: *Works and Worlds of Art*. Oxford: Clarendon.

## THE ASYMMETRY OF ‘BECAUSE’

Benjamin SCHNIEDER  
University of Hamburg

### *Summary*

The connective ‘because’ is often claimed to be asymmetrical. The paper examines (i) how to understand this claim, (ii) whether one can account for the asymmetry of ‘because’ in other terms, and (iii) how a number of *prima facie* counterexamples bear on it.

### §1. *Introduction\**

Compared to other natural language expressions, the connective ‘because’ has received surprisingly little systematic attention in the philosophy of language. Yet this connective deserves our attention since it plays a pivotal role in both everyday reasoning and scientific discourse (including philosophical discourse and reasoning as well).

This paper focuses on one alleged structural feature of ‘because’, namely its *directionality*, its being an *asymmetrical* connective. While the asymmetry of ‘because’ is often acknowledged and taken to be a characteristic feature of the semantics of ‘because’, there has been no focused discussion of the issue as of yet. This paper aims to close that gap: By delineating the

---

\* As a freshman in philosophy, I didn’t know anything about analytic philosophy. I wanted to understand the meaning of life and to study the deep thoughts of eminent figures such as Friedrich Nietzsche or Arthur Schopenhauer. As a finalist, I had already given up on the meaning of life; and on Nietzsche. Instead, I wanted to tackle delicate conceptual issues of a highly abstract character. My thesis dealt with Bolzano’s conception of substance and adherence, and what kept me awake were now questions such as: ‘What is the relationship between charm and Belmondo’s charm? Can Bolzano’s substances be characterised in terms of ontological independence? Can a substance have an essential attribute?’—What had happened to me!?

One culprit was Wolfgang Künne, who infected me with his interest in Bolzano. Another was Peter Simons. I had not yet met Peter back then. But his *Parts* featured in one of Künne’s seminars. I read the chapter on ontological dependence. It hurt my brain. Then I was hooked. And here I am, brain still hurting, still doing analytic metaphysics. Thank you, Peter.

important questions and surveying a range of possible answers, it prepares the grounds for a future debate. But it also takes a substantive stance on the questions (which can, however, for reasons of space only be defended in some depth).

Hence, I have three main objectives:

- To discuss how to precisely state the asymmetry of ‘because’; two different options will be shown to carry different commitments (§2).
- The asymmetry of ‘because’ is usually accepted on merely intuitive grounds. But can one do better? Can one explain *why* ‘because’ should behave asymmetrically? I will discuss three ideas about how the asymmetry of ‘because’, if accepted, can be accounted for in more fundamental terms (§3).
- Even though many philosophers take ‘because’ to behave asymmetrically, there are actually a number of examples in which a because-claim and its converse both seem acceptable. Consequently, I shall discuss whether and how one can defend the assumption of the asymmetry against these apparent counterexamples (§4).

## §2. *Stating the asymmetry of ‘because’*

Many philosophers share the intuition that because-sentences are internally directed and irreversible; for short, that ‘because’ is an asymmetrical connective. An early expression of this thought is Socrates’ central argument against Euthyphro’s analysis of piety. Let me briefly recapitulate the core of that argument (abstracting from exegetical issues and complications).<sup>1</sup> Euthyphro held that being pious consists in (or: is to be analysed as) being beloved by the gods. But if that analysis were correct, Socrates reasons, Euthyphro would have to accept the assumption that

A.1 The pious things are pious because the gods love them.

However, Euthyphro also holds (for independent reasons stemming from his theological views) that

A.2 The gods love the pious things because they are pious.

---

1. For the following take on the argument, cp. Sharvy (1972).

From A.2, Socrates infers:

C  $\neg$ (The pious things are pious because the gods love them).

But because of C, Euthyphro must reject A.1; hence, his analysis fails—or so the argument goes.

The inference from A.2 to C relies on the conviction that ‘because’ behaves asymmetrically: that the truth of a because-claim entails the falsity of the converse claim, which reverses the order of the main-clause and the because-clause. While Socrates does not put forward any explicit principle here, his way of arguing invites framing the asymmetry in terms of inferential properties of sentences: an operator may, for instance, always allow the sentences it connects to be permuted *salva veritate* (‘or’ is a case in question, while ‘if’ is not), such that any connectives possessing this property could be called *symmetrical*.<sup>2</sup> Following this idea, one can give a characterisation of the asymmetry of ‘because’ in terms of the validity of schemata:<sup>3</sup>

ASY<sub>SCHEMA</sub> If an instance of ‘ $p$  because  $q$ ’ is true, then the corresponding instance of ‘ $\neg(q$  because  $p)$ ’ is true as well.

Note how the asymmetry of ‘because’, thus understood, entails its irreflexivity:

IRR<sub>SCHEMA</sub> Any instance of ‘ $p$  because  $p$ ’ is false.

The current proposal of how to phrase the asymmetry of ‘because’ is semantically neutral: it is not committed to any particular underlying semantic framework which would specify a particular type of semantic value assigned to ‘because’ (or to sentential connectives in general). Unfortunately, as it stands, this proposal suffers from some imprecision given that natural languages allow for an abundance of ambiguity and context-sensitivity: instances of the schema ‘ $p$  because  $q$ ’ may allow for readings

---

2. Grice (1967, 67) characterizes the symmetry (‘commutativity’) of connectives in this way.

3. Philosophers who talk about structural properties of ‘because’ by using a similar approach include, e.g., Schnieder (2004a, 420), who takes ‘because’ to be asymmetrical, Lowe (1998, 145), who takes ‘because’ to be anti-symmetrical, staying agnostic about its irreflexivity; Künne (2003, 230) and Mulligan (2006, 38), who take ‘because’ to be irreflexive; Sharvy (1986, 513f.), who argues that Plato relied on the asymmetry of ‘because’ (or rather the Greek  $\tau\iota$ ).

with different truth-values affecting the validity of the schema as stated above. The issue will be discussed later (see §§4.a, 4.c).

An alternative way of stating the asymmetry of ‘because’ will suggest itself to many philosophers. As commonly understood, asymmetry is a property of relations:

$R$  is asymmetrical  $\leftrightarrow_{\text{df.}} \forall x \forall y (x \text{ stands in } R \text{ to } y \rightarrow \neg y \text{ stands in } R \text{ to } x)$ .

A straightforward way of stating the asymmetry of ‘because’ would therefore be in terms of the asymmetry of some relation *expressed* by the connective:<sup>4</sup>

ASY<sub>REL</sub> The connective ‘because’ expresses an asymmetrical relation.

Like the inferential characterisation of the asymmetry, the relational characterisation yields a corresponding characterisation of the irreflexivity of ‘because’ as an immediate corollary:

IRR<sub>REL</sub> The connective ‘because’ expresses an irreflexive relation.

Even if many philosophers find the relational characterisation natural, it should be acknowledged that it comes with a heavier theoretical burden than the inferential characterisation. The relational characterisation is

- (i) in need of additional information, and it is
- (ii) committed to a particular semantic framework.

*Re* (i): For the sake of completeness, the proposal must be supplemented with an explication of what it means that a connective *expresses* a relation, and with an account of what the *relata* of the relevant relation are. Both are controversial issues which cannot be decided here, though a few remarks should be helpful.

First, *expressing*—for present purposes, the semantic relation of expressing will be characterized in terms of truth-conditions: A connective  $C$  expresses a relation  $R$  iff the truth-conditions of a sentence  $S$  governed by  $C$  are stateable in terms of  $R$  holding between certain relata introduced by the clauses that  $C$  connects.<sup>5</sup>

4. See, e.g., Caputo (2007, 282), Bolzano (1837, §§ 168, 209), Correia (2005, 63), and Tatzel (2002, 17) (the latter three take ‘because’ to express an asymmetrical relation of grounding).

5. While the specified condition is arguably a minimal condition on any semantically useful

Second, the *relata*—if ‘because’ expresses a relation, we can derive constraints on the type of its *relata* from observations on admissible substitutions in because-sentences, since ‘because’ is a hyper-intensional connective: It is truth-conditionally sensitive to fine semantical differences in the clauses it connects. More specifically, it is sensitive to differences not straightforwardly accountable for in terms of their purely modal behaviour (i.e. to differences between clauses that are co-equivalent with respect to all possible worlds).<sup>6</sup> This is illustrated by Aristotle’s famous insight that the first of the following sentences is true while the second is false:<sup>7</sup>

AI That snow is white is true because snow is white.

AI<sub>conv</sub> Snow is white because it is true that snow is white.

These sentences differ in truth-value even though their clauses are modally equivalent (true in exactly the same possible worlds). So, ‘because’ takes into account differences that outrun purely modal differences.

A direct consequence follows for the *relata* of the relation expressed by ‘because’: Such *relata* must be subject to a considerably fine-grained individuation—they cannot, for instance, have the coarse grain of sets of possible worlds (or Lewisian propositions).<sup>8</sup> Suitable *relata* might be structured Bolzanian propositions,<sup>9</sup> or Finean (1982) F-Cond facts, though even something more fine-grained might be needed.<sup>10</sup> The issue cannot be decided here, but we need some vocabulary to move on; by *terminological fiat*, the pertinent *relata* will be called the *contents* of the clauses in the because-sentence, or simply *propositions*.

---

notion of *expressing*, it may not be sufficient for certain purposes. Depending on how relations are individuated, the condition may for instance allow that a connective expresses a multitude of relations, some of which might seem to import semantically irrelevant information. If, for instance, the relation of *happening later than* and the relation of *being a round square or happening later than* are different, then the second one might look like a worse candidate for being expressed by the connective ‘after’ than the first one. But this paper is not the place to explore such issues.

6. Cp. Schnieder (2011), which develops, starting from this observation, a logical system for ‘because’ in which even logically equivalent formulas cannot be substituted *salva veritate* in the scope of ‘because’.

7. See *Metaphysics*, book Θ 10: 1051<sup>b</sup>6–9. The question of how to account for Aristotle’s insight plays an important role in the contemporary debate about truth; see, e.g., Künne (2003, 150–57), Hornsby (2005, 43f.), and Rodriguez-Pereyra (2005).

8. Cp. Ruben (1990, ch. 5), Schnieder (2010, 324f.), or Schnieder & Steinberg (forthcoming).

9. On such a conception, see Künne (2003, 42–52, *et passim*).

10. See Schnieder (2010, especially 337–40).

*Re* (ii): As we have seen, the relational characterisation of the asymmetry of ‘because’ needs supplementary information; but there is another issue worth mentioning. Unlike the inferential characterisation, the relational account is *not* semantically neutral. Instead, it seems to incur a substantial presupposition about the semantic function of ‘because’ and, presumably, of connectives in general: namely that they are in the semantic business of expressing relations.

Of course, we are accustomed to treating connectives that way; the treatment is indeed part of semantic frameworks in the spirit of Montague, which are widely used in philosophy and in linguistics. Nevertheless, it should be admitted that assigning relations to connectives was not an *obvious* choice, for terms that express relations are, in the first instance relational predicates which accept singular terms as arguments, and sentential connectives are no such thing, since they do not attach to singular terms. So one wonders why someone using them should (inadvertently) commit himself to there being some expressed relations, as well as an appropriate domain of relata between which those relations hold.<sup>11</sup> While it has nevertheless proven convenient to use relations as the semantic values of connectives, it is not clear that this treatment is faithful to the actual semantic purpose of connectives. An alternative of how to treat connectives in a semantic theory might, for instance, state semantic clauses for them with the aid of *connectives* belonging to the meta-language.<sup>12</sup>

So far, nothing I have said is meant to argue that the relational characterisation of the asymmetry of ‘because’ must be defective; the goal was only to show that it comes with some theoretical burdens—burdens that may be perfectly acceptable in the end. And even for those who take these worries seriously, the relational characterisation may remain a helpful and workable tool. So, for present purposes, I will not commit myself to the superiority of either characterisation. Instead, by working with both, I aim to identify some of the implications picking one has on the matters discussed.

### §3. *Accounting for the asymmetry of ‘because’: three proposals*

As we have seen, many philosophers accept that ‘because’ is an asymmetrical connective. If one examines their reasons, it appears that most endorse

---

11. For worries of roughly this sort, see, e.g., Hornsby (2005 42f.) and Liggins (2008, 189).

12. Compare also Williamson’s idea (1999, 259ff.) that a semantic account of quantification into sentential position should not reduce it to nominal quantification in the meta-language.

the thesis by virtue of intuitions arising from their general understanding of ‘because’, or perhaps on the basis of an inductive generalization from examples of irreversible because-claims. Such examples abound in both philosophical contexts (such as the cases cited from Socrates and Aristotle) and ordinary ones. In particular, because-claims used to report causal relations between individual events seem to support the thesis: Someone who claims that the Titanic sank because it hit the iceberg should deny that the Titanic hit the iceberg because it sank.

While believing in the asymmetry of ‘because’ on an intuitive and/or inductive basis is not necessarily objectionable, it would be nice if one could rest her conviction of the asymmetry of ‘because’ on firmer grounds. In other words, an *argument* for the asymmetry would be preferable. Even nicer than a mere argument to the effect *that* ‘because’ behaves asymmetrically, however, would be an explanatorily virtuous argument which also provides an account of *why* ‘because’ is asymmetrical. Rooting the asymmetry in some underlying, more fundamental facts, would certainly deepen our understanding of it (though one must acknowledge the possibility that the asymmetry of ‘because’ holds only as brute fact).

To my knowledge, the issues raised here have not been addressed so explicitly before. Nevertheless, some proposals of how to account for the asymmetry of ‘because’ can be distilled from the existing literature. I will go through the three cases that seem most important to me.

a. *The asymmetry of explanation*

One might want to account for the asymmetry of ‘because’ in terms of what is called the *asymmetry of explanation*.<sup>13</sup> Whether (and how) such an account could work depends on the meaning of ‘explanation’, which is a highly flexible and ambiguous term. Being derived from the verb ‘explain’, the term denotes in its central meaning a sort of *communicative act*; this is the meaning I adopt in this section.

Sometimes, however, philosophers mean by ‘explanation’ either the *linguistic vehicle* used to perform such an act—i.e. a sentence—or the *semantic content* of such a vehicle. Accordingly, what those philosophers mean by the *asymmetry of explanation* is a claim about sentences or about the contents of sentences, i.e. propositions. In the first case, the asymmetry

---

13. The asymmetry of explanation plays an important role in the literature on explanation, and it is widely (though not universally) endorsed: see, e.g., Nozick (1981, 116), Horwich (1987, 155), Newton-Smith (2000, 128), or Bird (2003, 301).

of explanation just *is* the asymmetry of ‘because’ and we cannot account for one in terms of the “other”. If ‘explanation’, however, is used to denote the contents of because-sentences, the asymmetry of explanation indeed seems the more fundamental phenomenon than that of ‘because’. For the latter is certainly not a brute accident of the English language (e.g. resulting from an arbitrary grammatical convention) but results from what the sentences *mean*—after all, synonyms of ‘because’ in other languages exhibit the same asymmetry. This would be a highly unlikely and inexplicable coincidence if the asymmetry of ‘because’ did not derive from its meaning. It therefore seems that the asymmetry of because-sentences depends on an asymmetry in the contents they express. However, this observation does not provide any substantial account of the asymmetry of ‘because’ yet; it merely stresses that the asymmetry is a semantic phenomenon and thus puts plausible constraints on where any account of it must start.

Things are different if one uses ‘explanation’ in its primary sense as applying to explanatory acts. Then one can propose a substantial account of the asymmetry of ‘because’ in terms of the asymmetry of explanation: Since (i) explanations behave asymmetrically and since (ii) because-sentences are used to give explanations, those sentences inherit the asymmetry of explanation—or so the argument would go. Deepening our understanding of the asymmetry of ‘because’ would then require directing our attention to the nature of explanatory acts, and to why they involve some sort of asymmetry. One idea might be that the epistemic dimension of explanations is responsible for their exhibiting an asymmetrical behaviour. Roughly put: Information used to explain something influences the epistemic system of the tutee (the person in need of an explanation), resulting in certain epistemic changes: Maybe an increased *understanding* of the relevant phenomena,<sup>14</sup> or perhaps changes in *credences*.<sup>15</sup>

That idea clearly needs work. But I have doubts about any approach starting from the thought that the meaning of ‘because’ should be understood from its use as a vehicle for explanatory speech acts. Such thought looks suspiciously like an instance of what Searle (1969, 136–40) dubbed the *speech act fallacy*. As a general rule, Searle held, we should not try to derive semantic insights about expressions from the simple fact that they are used to perform certain speech acts. The properties of speech

---

14. For the common idea that explanation is intimately linked to understanding see, e.g., Bromberger (1965) and Achinstein (1983), but also critically Trout (2002).

15. Cp. Gärdenfohrs (1980); a link to asymmetry is suggested in Gärdenfohrs (1990, 118).

acts are subject to a manifold of pragmatic aspects without any straightforward implications about the semantics of the expressions used to perform them.

Still, one may accept Searle's point as a general maxim and yet differ about particular cases. There might be certain expressions so intimately connected to the performance of a particular speech act that their semantics are after all constrained by the properties of that speech act. The case at hand might be an example. To some philosophers, 'because' may seem to be inextricably connected to explanation—an explanation just *is* the utterance of a because-sentence and vice versa. Friends of the popular slogan that why-questions are requests for explanations may well think this way (van Fraassen 1980, 134); after all, a because-sentence is what directly answers a why-question.

But, as pointed out already in the early literature of explanation, the identification of explanations and utterances of because-sentences seems linguistically shortsighted.<sup>16</sup> Firstly, apart from explanations of *why* something is the case, there are other sorts of explanation, e.g. an explanation of *how* something was done, or of *what* the relationship between certain facts is, or even of *who* you are. Explanatory speech acts can be performed by a variety of linguistic vehicles apart from because-sentences.

Secondly, utterances of because-sentences need not be explanations, for it is important to distinguish between *merely telling* someone why *p* and *explaining* to someone why *p*. An utterance of a because-sentence may fail to be explanatory because it does not involve sufficient information to enlighten the addressed tutee—even though he has been *told* why *p*. Here is an example: Fred knows that *p*. And for all he knows, the fact that *p* must obtain for one of two grounds, that *q* or that *q*\*. Unfortunately, his epistemic background also gives him good reasons to exclude both possibilities. He is in an epistemic predicament and what is called for is an explanation of why *p*. Assume now that Ann, who Fred accepts as an authority in the matter, informs him that *p* because *q*. But she does this without elucidating the matter for him any further; she merely states the bare fact. Since Fred finds Ann trustworthy, he may well believe her and accept that *p* because *q*. But yet, the matter remains obscure to him as before. He has been (correctly) *told* why it is the case that *p*, but he has not been *explained* why it is the case that *p*. Nothing has been done to solve his epistemic predicament, and nothing has been done to enable him to

---

16. On the following two points, see Bromberger's seminal paper (1965).

*understand* why it is the case that *p*. To explain to Fred why *p* (and not just tell him why *p*), one would have to provide him with information that helps remove his reasons which spoke against the hypothesis that *p* because *q*.

Given the distinction between telling and explaining why *p*, it seems that the *primary* speech act performed by uttering because-sentences is not that of *explaining*, but that of *telling* or *informing* why something is the case. Whether an act of *telling why* additionally counts as an act of *explaining why* depends on further conditions on the communicative context (in particular on the epistemic situation of the tutee). Inasmuch as such further conditions determine whether an act of *telling how* something was done also counts as an act of *explaining how* it was done.

For the reasons given, we ought to sharply distinguish a theory of explanation from the project of stating the semantic rules that govern our use of 'because'. Approaching the connective 'because' via reflections on explanatory acts or intuitions on the use of 'explain' and 'explanation' invites confusion over its semantics. The prospects of accounting for the asymmetry of 'because' in terms of the asymmetry of explanation seem dim in that light, and I will not pursue that approach any further here.

b. *The asymmetry of 'because' and other structural properties*

If a relation is irreflexive and transitive, it must be asymmetrical. Lange (2009, 207) relies on this fact in arguing for the asymmetry of explanation: Explanation is irreflexive and transitive. Therefore, it is asymmetrical. The argument, if acceptable, can easily be tailored to suit present purposes: 'because' is an irreflexive and transitive connective, hence it is asymmetrical.

A dialectical problem of (both versions of) the argument is that it justifies a controversial claim in terms of a claim that is at least as contentious, since philosophers have voiced serious doubts about the transitivity of explanation and 'because'. Standard arguments against transitivity point out that chaining true because-sentences can easily lead to unacceptable ones (Hesslow 1981; Owen 1992, 15–19). A stock example is provided by a nursery rhyme: 'For want of a nail the shoe was lost, for want of a shoe the horse was lost, for want of a horse the rider was lost, for want of a rider the battle was lost, for want of a battle the Kingdom was lost, and all for the want of a horseshoe nail.' Each of the lines corresponds to a

true because-statement ('the horse lost its shoe because a nail was missing,' etc.). But the result of chaining them looks bizarre:

MIGHTY-NAIL The Kingdom was lost because a nail was missing.

Hence, the argument continues, 'because' is not transitive.

Whether the argument establishes its goal is certainly disputable. The truth-conditions of because-sentences have to be distinguished from the acceptability-conditions of utterances of because-sentences. That MIGHTY-NAIL looks bizarre need not indicate that it is *false*. The sentence might rather, in normal circumstances, be inapt for successfully performing an explanatory act. Relatedly, although a because-sentence can clearly be true even if its because-clause provides only a partial and/or remote ground of its main clause, uttering '*p* because *q*' usually signals, due to conversational principles, that the because-clause provides a salient reason for the truth of the main clause. Citing highly partial and remote reasons will therefore normally result in an unacceptable utterance. In fact, the story of the rhyme clearly warrants a because-sentence which explicitly stresses the partiality of the given explanation:

NOT-SO-MIGHTY-NAIL The Kingdom was lost because, *among other things*, a nail was missing.

If one thinks that *semantically* 'because' allows for as much partiality as 'because among other things', one should not reject MIGHTY-NAIL as false, although it may be pragmatically defective.

But there are other apparent failures of transitivity which may seem more robust.<sup>17</sup> Whatever stance we take on the transitivity of 'because', though, the fact that it is highly controversial makes it an unlikely candidate to lend support to the asymmetry of 'because'.

Moreover, even if one thought that we have a (perhaps weak) argument for asymmetry here, the argument can hardly be seen as *accounting* for the asymmetry, i.e. as providing insight into underlying facts which *make* 'because' asymmetric.

---

17. See, e.g., Putnam (1973) and Schaffer (2012); see Litland (2013) for a reply to Schaffer.

c. *The asymmetry of 'because': An Aristotelian account*

A significant (though indirect) contribution to a theory of 'because' is found in Aristotle's works. Two central elements of the Aristotelian view are: (i) If we want to know *why* something is the case, we want to find out about certain objective priorities. (ii) Such priorities can differ in sort; they can, e.g., be causal, they can concern what things consist of, or they can concern the essences of things.

Since because-sentences report why something is the case, the Aristotelian picture delivers a partial account of the semantic function of 'because': the connective tracks relations of objective priority. A full Aristotelian semantics of 'because' should address several further issues: first, what is a priority relation and is there anything distinctive of the priority relations tracked by 'because' (clearly, not any old priority relation supports a because-sentence: mere temporal priority is a clear case in question)? And second, how exactly are the clauses of a because-sentence related to the relata of the supporting priority-relation?

But acquaintance with the basic idea of the semantics suffices to capture the relevant point: once an Aristotelian semantics of 'because' is in play, it provides a straightforward account of the asymmetry of 'because'. Priority relations are asymmetrical by their nature, and since the meaning of 'because' is to track priority relations, those relations bestow 'because' with its directional feature of asymmetry.<sup>18</sup>

Admittedly, this account of the asymmetry of 'because' is not particularly *deep*. The phenomenon explained (the asymmetry of 'because') is still very close to what explains it (the asymmetry of priority relations together with the thesis that 'because' tracks priorities). Or is it even identical to it, such that we have been given a pseudo-explanation, a mere relabelling of a phenomenon?

No. To see this, consider  $ASY_{REL}$  which frames the asymmetry of 'because' in terms of an expressed asymmetrical relation. Given an Aristotelian semantics, one might be tempted to take 'because' to *express* the priority relations it tracks. That 'because' expresses an *asymmetrical* relation would then be accounted for by its expressing a *priority* relation, which is not the same thing (*being taller than* and unrequited love are asymmetrical relations, but not priority relations). But note that one need not identify the relation expressed by 'because' with the priorities it tracks. Instead, an

---

18. Cp. Aristotle's *Anal. Post.* Bk. I, part 3.

Aristotelian semantics may take ‘because’ to express a higher-order relation obtaining between two propositions iff there is a priority relation holding between relata introduced by those propositions. This lends ‘because’ a univocal meaning; moreover, although the higher-order relation expressed by ‘because’ has propositional relata, it can hold in virtue of priority relations holding between relata of other categories (e.g. events, things and essences, concepts, etc.). On this account, the propositional relation expressed by ‘because’ inherits its asymmetry from the priority relations that base the higher-order relation.

Some philosophers have flirted with the idea that explanatory asymmetry is owed to the asymmetry of causation. As a general account of the asymmetry of explanation the idea clearly fails, because asymmetry also occurs in non-causal explanations.<sup>19</sup> However, on the current picture, the idea is correct as a *partial* account of the asymmetry of ‘because’ (and if one means by an explanation a because-claim, also as a partial account of the asymmetry of explanation).<sup>20</sup> Hence, one priority relation supporting ‘because’ is indeed causation.

Another topic noteworthy in this context is *metaphysical grounding*. In the current debate, grounding is taken to be a relation structuring the facts of reality, so that less fundamental facts depend on (or: are grounded in) more fundamental facts (which are the grounds of the former). Moreover, grounding is usually taken to be an asymmetrical relation.<sup>21</sup> Finally, it is commonly assumed that the notion of grounding can be introduced by the aid of because-claims: If the fact that *p* is grounded in the fact that *q*, then the former obtains because the latter does. However, most proponents of metaphysical grounding reject the reverse claim, as they take grounding to be different from causation, and a because-sentence can be causal. Now, given an Aristotelian semantics of ‘because’, there is a straightforward account of the connection between grounding and ‘because’: metaphysical grounding is one of those priority relations supporting because-claims (in fact, grounding is sometimes described as something like metaphysical causation) (Schaffer 2012, 122). The asymmetry of grounding would then partially be responsible for the asymmetry of ‘because’; not *solely* responsible, though, since ‘because’ tracks other priority relations as well, in particular causation.

---

19. As Hausman (1982: 43) pointed out following Railton (1980).

20. Hausman (1993) also holds that the idea is correct for causal explanation.

21. See, e.g., Rosen (2010: 115), Correia & Schnieder (2012), or Fine (2012).

#### §4. *Counterexamples and defences*

Even though many philosophers take ‘because’ to be asymmetrical, there are apparent counterexamples in the form of true because-sentences where the order of the clauses can apparently be reversed *salva veritate*. In what follows, I discuss a number of different classes of such examples. Some of them are easily recognizable as merely *superficial* counterexamples; however, a discussion of them may nevertheless be helpful for revealing resources one can employ in the face of more threatening cases.

##### a. *Ambiguities*

For a start, consider two simple cases: A financial institute situated on a riverbank was blown up with loads of dynamite, such that the explosion also destroyed the riverbank. In that context, the following sentence seems to be true:

- (1) The bank was destroyed because the bank was destroyed.

Or take a slightly modified case: A riverbank was blown up with loads of dynamite, such that the explosion also destroyed the riverbank on the opposite shore. Then it is true that:

- (2) *That* riverbank was destroyed because *that* riverbank was destroyed.

The two sentences are superficial violations of the irreflexivity of ‘because’, and hence of its asymmetry. But they are *only* superficial problems. Clearly they do not violate the asymmetry of ‘because’ in its relational characterisation. While (1) has a reading in which it is true, it is a reading in which the two clauses of (1) do not express the same proposition. So, although (1) contains the same clause twice over, in the relevant reading it will deliver different relata for the relation expressed by ‘because’—in a reading where it delivers the same relatum twice over, however, the sentence is false in the specified context (it is not that the financial institute was destroyed because *it* was destroyed). The same holds for (2), the only difference being what accounts for why two different relata are delivered: in the case of (1), it is due to a semantic ambiguity of ‘bank’, while in the case of (2) it is due to the context-dependence of the demonstrative ‘that’.

The examples show, however, that the asymmetry thesis phrased in terms of schemata fails as a general claim about sentences with a certain *surface structure*. This is no surprise. Universal claims about the truth or falsity of natural language sentences with a particular surface structure are generally liable to be false because of features like ambiguity and context-dependence.<sup>22</sup> So if the asymmetry thesis is phrased in terms of schemata, a precise rendering requires some specifications of what counts as admissible instances of the schemata: We need to exclude different readings of ambiguous expressions in the reversed because-sentences and different indexical referents.

Turning to a related example, Kurt Tucholsky once made the following sociological remark, which might well be an empirical truth:

- (3) Most people celebrate Christmas because most people celebrate Christmas.

Here again we have a failure of irreflexivity and asymmetry, but only a superficial one: The sentence has an ambiguous syntax, allowing for two readings represented by the following semi-formalisms:

- (3.1) (For most  $x$ :  $x$  celebrates Christmas) because (for most  $x$ :  $x$  celebrates Christmas).  
(3.2) For most  $x$ : ( $x$  celebrates Christmas because for most  $y$ :  $y$  celebrates Christmas).

That (3) may be true is due to the second reading, which is true if sufficiently many people satisfy the open sentence ‘ $x$  celebrates Christmas because for most  $y$ :  $y$  celebrates Christmas’; and there may be enough people who do that. The first reading of (3), however, does not even appear to be potentially true.

How does the ambiguity of (3) bear on the asymmetry thesis? Reading (3.1) can be ignored since it delivers a false statement. That (3) allows for reading (3.2) does not violate the asymmetry thesis phrased in terms of a relation between propositions. The truth of (3.2) would turn on the question whether the explanatory relation holds between sufficiently many singular propositions (of the form:  $x$  celebrates Christmas) and a quantified proposition; no explanatory relation between one and the same proposition is involved.

---

22. Cp. Strawson (1957: 19ff.).

But the example makes another modification necessary if one works with the inferential characterisation of the asymmetry, because merely banning different resolutions of semantic ambiguities and indexical elements does not suffice to give us a valid schema. The formulation can be improved by referring to the logical form of the instances of the schemata: Only those instances are relevant for which ‘because’ has widest scope on the level of logical form. Alternatively, one could demand the two instances of ‘ $p$ ’ express the same proposition (equally for ‘ $q$ ’). For in readings in which a quantifier has widest scope, the relevant instance of ‘ $p$ ’ will not express a proposition but rather introduce a propositional fragment which partly provides the quantificational structure of the proposition expressed by the whole sentence. One could then cover all restrictions (constancy of linguistic meaning, constancy of reference, a particular logical form) in one condition:

ASY<sub>SCHEMA\*</sub> If an instance of ‘ $p$  because  $q$ ’ is true, so is the corresponding instance of ‘ $\neg(q$  because  $p)$ ’, if both instances of ‘ $p$ ’ and ‘ $q$ ’ express the same proposition respectively.

Note that the logical form of a because-statement may depend on *hidden* constituents. Assume Jo went to a dull magician’s show because she was invited; since she knew the magician, she definitely expected no surprises. But the magician outgrew himself and managed to surprise Jo. Realizing that her expectations were wrong, she was surprised again. Hence:

(4) Jo was surprised because Jo was surprised.

The example violates the asymmetry of ‘because’ on the surface level, but arguably not on the level of logical form. The predicate ‘is surprised’ receives a temporal argument, which, if left unspecified, is provided by an implicit (suitably restricted) existential quantifier. A reading which renders (4) true will assign the following form to it:

(4.1)  $\exists t \exists t^*$  (Jo was surprised at  $t$  because Jo was surprised at  $t^*$ ).

(The first variable might be restricted to times preceding the utterance, the second to times preceding the value of the first variable. Details are irrelevant here.)

But (4.1) itself does not have the right form to directly violate the asymmetry thesis. It would still cause a problem for the thesis if (4.1) was

made true by an interpretation in which the same value is assigned to both variables—but such an assignment does not make it true.

Example (4) is instructive for a whole class of cases that might initially seem problematic for the asymmetry thesis, namely cases of causal feedbacks. In certain scenarios, we may agree that *F*s occur because *G*s occur *and vice versa*; for instance, sometimes uproars arise because military interventions take place and vice versa; similarly, unemployment can increase because profits decrease, and vice versa. Do such examples violate the asymmetry of ‘because’ and necessitate some further restriction (perhaps to explanations of singular events, whereas explanations of types of events are excluded from the scope of the thesis)?<sup>23</sup> No. The distinctions made above suffice for seeing that such examples are compatible with the asymmetry of ‘because’. The claim that, in a reciprocal system, *F*s occur because *G*s occur and vice versa, involves some sort of quantification; for the sake of simplicity, let us treat it as universal quantification. Making explicit its logical form reveals, then, that no violation of the asymmetry of ‘because’ is involved, since its form will roughly be:

$$\forall x(Fx \rightarrow \exists y(Gy \ \& \ x \text{ occurs because } y \text{ occurs})) \ \& \\ \forall y(Gy \rightarrow \exists x(Fx \ \& \ y \text{ occurs because } x \text{ occurs})).$$

(Depending on the specific claim, temporal indices may also play a role; moreover, instead of a universal quantification, an existential or some sort of generic quantification can be relevant; it is unnecessary to go into more detail here.) Only if a single pair of values satisfied ‘*x* occurs because *y* occurs’ and equally ‘*y* occurs because *x* occurs’, would we have a violation of asymmetry. But that would involve mutual causation between two particular events, and reciprocal systems do not force us to allow for such cases.<sup>24</sup>

#### b. *Implicature*

Let us turn to a different sort of example. In *The Two Gentlemen of Verona* (Act 1, Scene 2), Julia questions Lucetta why she thinks Proteus is the best. Lucetta replies:

---

23. Michael Baumgartner suggested to me that such a restriction is needed.

24. See, e.g., Frankel (1986).

- (5) I think him so because I think him so.

Again we have an apparent violation of the irreflexivity, and hence asymmetry, of 'because'. Yet it seems clear that the counterexample is spurious. After all, almost everyone will straightforwardly interpret Lucetta as conveying something different, namely that she *lacks* a reason for her thought. On this interpretation, she does not literally mean what she said; rather what she said is used to generate a conversational implicature with a different content. By speaking in such cases as Lucetta's of the 'because' of the exasperated adult, Mulligan (2006: 38) draws attention to the fact that even children can easily understand this sort of implicature.

So the example clearly provides no argument against the irreflexivity of 'because' (or, thereby, against the asymmetry). To the contrary: That we immediately understand Lucetta as communicating not that she has a very strange belief about explanatory relations, but rather as communicating that she lacks a reason for thinking Proteus to be the best, is telling and can even be used in a small argument *in favour* of the irreflexivity of 'because'. For the implicature to get off the ground it seems essential that we regard the sentence uttered by Lucetta as obviously false. Only because we reject instances of '*p* because *p*' straightaway in their literal meaning are we warranted in inferring that Lucetta must be conveying something else (compare the mechanism of irony). Hence, our readiness to interpret the 'because' of the exasperated adult as communicating a lack of reason is at least an indication that we accept the irreflexivity of 'because'.

c. *The evidential 'because'*

Alice: Oh, it's getting warmer. Queen: Why? Alice: Because the thermometer is rising. The reply Alice gave is elliptical, and its spelled out version would read:

- (6) It's getting warmer, because the thermometer is rising.

While Alice made an adequate reply, we also know that (6) runs contrary to the causal facts. And we can report the causal facts by the true sentence:

- (6\*) The thermometer is rising because it's getting warmer.

Sentences (6) and (6\*) are an apparent counterexample to the asymmetry of ‘because’.

It is indisputable that (6) and (6\*) can both be acceptable with respect to the same thermometer. Nevertheless, the asymmetry can be upheld. To prepare for the argument, let us focus on sentence (6). It illustrates a peculiar use of because-sentences in which they do not serve to state why the main clause is true; instead, the because-clause specifies *evidence* for the truth of the main clause (i.e. a reason for believing it). Call this the purely *evidential* use of ‘because’.

The evidential use of ‘because’ has been well-examined in linguistics. It is distinguished from the non-evidential use by a number of linguistic markers: (i) It resists cleft constructions, ‘because of’-transformations, and straightforward negation: ‘That *q* is why *p*’, ‘*p* because of *x*’, and ‘Not *p* because *q*’ trigger a non-evidential reading. To negate ‘*p* because *q*’ in an evidential reading, you use ‘That *q* doesn’t mean that *p*’.<sup>25</sup> (ii) It often has some epistemic modal in its main clause, and has a lower degree of acceptability otherwise.<sup>26</sup> (iii) It favours the use of a comma (whereas the non-evidential fares better without one).<sup>27</sup> (iv) It requires a longer computation time for interpretation, particularly if it occurs without any epistemic modals.<sup>28</sup>

Together, these observations provide good reason to regard ‘because’ as being semantically *ambiguous* between an evidential and a non-evidential reading.<sup>29</sup> And even if they do not make this explicit, it seems clear that those who take ‘because’ to be asymmetrical intend to speak only about the non-evidential uses of ‘because’ (also the Aristotelian semantics then is meant to apply to the non-evidential ‘because’ only). So the asymmetry thesis (be it in the inferential or the relational characterisation) should henceforth be restricted to the non-evidential ‘because’. This is in no way an *ad hoc* or artificial restriction; someone who makes a claim about the semantic properties of some other ambiguous term, such as ‘bank’ or of ‘healthy’, will in general equally restrict it to one of the semantic senses of the term.

We can then see that the asymmetry *in question* is unthreatened by examples such as (6) and (6\*). What they show is that an evidential

---

25. Cp. Bender & Kathol (2001).

26. Cp. Morreal (1979: 236).

27. Cp. Kac (1972).

28. Cp. Mohammed & Clifton (2008).

29. Although the two uses are plausibly related their connection need not be discussed here.

because-statement can be true while the reverse non-evidential statement is also true. But this does not conflict with the asymmetry that is at stake here, i.e. the asymmetry of 'because' in its non-evidential sense. (Note in passing that the evidential 'because' itself does not seem to be asymmetrical, since two clauses can presumably mutually support each other in an evidential fashion; but again, this does not bear on the asymmetry of the non-evidential sense of 'because'.)

d. *Priorities in harmony?*

The examples we have focused on so far were no genuine threat to the asymmetry of 'because'. They should make us aware, however, that the proper evaluation of a because-claim often turns on a number of contextual clues; this observation is useful for defending the asymmetry of 'because' from more serious challenges. In particular, the evidential/non-evidential ambiguity of 'because' can be important here. Since that ambiguity seems to pervade our language, it need not always be obvious whether, when we find a because-statement acceptable, we rely on an evidential or on a non-evidential reading.

Now let us turn to a more philosophically challenging case. Recall the following step in Socrates' argument against Euthyphro's analysis of the pious:

- A.2 The gods love the pious things because they are pious.
- C  $\neg$ (The pious things are pious because the gods love them).

But this conclusion contradicts the reverse claim

- A.1 The pious things are pious because the gods love them.

whereas Euthyphro's analysis of the pious as the god-beloved would commit him to A.1. Hence he should abandon his analysis (given that he does not want to give up A.2, which flows from his conception of the gods).

The argument has been criticised by a number of commentators for its reliance on the asymmetry of 'because'. The alleged problem is that there are different *senses* of 'because', e.g. a causal and a non-causal one, while the asymmetry must be restricted to a single sense.<sup>30</sup> In other words, the

---

30. The criticism is due to Brown (1964). His claim that 'because' is ambiguous has been explicitly accepted by several commentators; see, e.g. Hall (1968: 6), Cohen (1971: 7), and Paxson (1972: 178–80) (though they argue that the ambiguity is not fatal for the argument).

move from ' $p$  because  $q$ ' to ' $\neg(q$  because  $p)$ ' is only warranted if the sense of 'because' is held constant. In A.2, 'because' has the sense of a rationalization, and so C is justified in that very sense of 'because'.<sup>31</sup> But in A.1, 'because' is used in a different sense, call it the *constitutive* one. So, both A.1 and C can be true at the same time, since they cannot be shown to contradict each other by the asymmetry of 'because'.

As it stands, the objection presupposes that 'because' has different senses. This presupposition is typically justified by pointing out that different sorts of explanation can be given with the use of 'because'; specifically, we can give a rationalizing explanation, as exemplified by A.2, and a constitutive one, as exemplified by A.1. In itself, however, that hardly shows that 'because' is semantically ambiguous. The Aristotelian semantics suggested earlier can easily account for the data without positing an ambiguity, if only it declares that because-statements can be based on different priority relations.

Regardless, assume we buy the ambiguity of 'because' for the sake of the argument. What is shown, then, is *not* that the asymmetry of 'because' breaks down when different senses of 'because' are involved. At best we should conclude that we cannot simply *assume* that asymmetry applies even if different senses of 'because' are involved. In the case of the Euthyphro argument, this means that we cannot justify the move from A.2 to C where the latter involves the constitutive sense of 'because' *merely* on the conviction that 'because' is asymmetrical in its single senses. We would need additional justification for that move.

To pinpoint what is at stake here, I will call two priority-relations  $R$  and  $R^*$  harmonious iff whenever  $xRy$ , it is not the case that  $yR^*x$ . And I will call 'because' *harmonious* iff it is not associated with any non-harmonious priority relations  $R$  and  $R^*$ . Talk about *associated* priority relations is meant to be neutral about the question of whether the non-evidential 'because' is semantically ambiguous or not; depending on whether it is or not, a precise statement of harmony would differ slightly: If 'because' is ambiguous, we can call 'because' harmonious iff whenever ' $p$  because  $q$ ' holds in one (non-evidential) sense, ' $q$  because  $p$ ' does not hold in any other (non-evidential) sense of 'because'. If 'because' is univocal but can be based on different sorts of priority relations, we can call it harmonious iff all priority relations that can ground because-statements are harmonious.

---

31. C would even seem trivially true in the rationalization sense of 'because', since the truth of ' $p$  because  $q$ ' in that sense requires ' $p$ ' to specify actions or attitudes of sentient beings for which the 'that'-clause then provides a reason.

Returning to the asymmetry of ‘because’, what has to be decided is the *Question of Harmony*: is ‘because’ harmonious or not? If it is harmonious, then the asymmetry thesis requires no revision, even though ‘because’ can be associated with different priority relations. But if ‘because’ is associated with some disharmonious priority relations, asymmetry has to be restricted accordingly.

Unfortunately, I know of no general argument that could decide the question of harmony. But I will consider three cases from the literature that allegedly illustrate non-harmonious sorts of priorities.

The first case is due to Anthony Kenny (1998, 27). Imagine a judge called Othon. If we ask ourselves why Othon is a judge, we see that the following is true:

(7) Othon is a judge because he judges.

But now let us ask: why does Othon judge? We then see that the following holds:

(7<sub>R</sub>) Othon judges because he is a judge.

We have a violation of unrestricted asymmetry here because two disharmonious priority relations are associated with ‘because’ in (7) and (7<sub>R</sub>) respectively—or so Kenny’s argument goes, reformulated in our current terminology.

Let us take a closer look at the two because-claims involved, starting with (7<sub>R</sub>). The idea is that Othon habitually judges because he is a judge, i.e., because he has a certain profession and having that profession makes him perform certain duties. That sounds like a correct explanation, arguably of a causal nature (it can certainly be enriched by additional information, e.g. that Othon is able to perform his duties and is willing to do so, but such explanations can be correct and partial).

What about the other direction, i.e. (7)? Kenny regards the ‘because’ here as expressing a constitutive explanation: that Othon judges non-causally makes him a judge, hence he is a judge because he judges. The statement may sound plausible at first, but one soon finds reasonable doubts. What exactly is meant by the verb ‘judge’ here? Presumably, it means (following the *OED*) to pronounce sentence upon (a person) in a court of justice. But then, to judge is an action only performable by people in a certain office, namely by judges. Being made a judge (by

some lawful ceremony, or in whatever way is suitable according to the relevant legal system) is what enables someone to judge (in the pertinent sense). And hence, the constitutive relation that Kenny sees does not exist: It is not because someone judges that he is a judge. That sentence (7) sounds acceptable at first is likely due to an evidential reading of it—A: He's a judge. Q: Why (implicit addition: do you say/think so)? A: Because he judges.—But in that reading, (7) is irrelevant to the asymmetry thesis.

Second case: the Euthyphro-argument itself. The objection to it described above must assume that the relevant priority relations associated with 'because'—rationalizing priority, which is associated with A.1, and constitutive priority, which is associated with A.2—are non-harmonious. To properly evaluate this claim, one needs at least some account of the rationalizing use of 'because'. What seems crucial to this use is that if ' $x$   $\phi$ -s because  $p$ ' is true, then the fact that  $p$  gives  $x$  a reason for  $\phi$ -ing. This, in turn, presupposes that the fact that  $p$  is epistemically accessible to  $x$ , and that  $x$  indeed accesses it or responds to it. Hence, the fact is there to be so accessed. But then, the fact must obtain independently of  $x$ 's  $\phi$ -ing. If ' $p$  because  $q$ ' is true in the constitutive sense of 'because', however, the fact that  $p$  does not obtain independently of the fact that  $q$ . Therefore, the disharmony of the two senses is doubtful.

A note in passing: at first glance, what has been said may seem to conflict with the idea that there can be *response-dependent* concepts or predicates which apply to a thing (partially or even wholly) because certain people think they apply to them—for example, it may be plausible that what makes someone cool is (ultimately) that people think he is cool. But the asymmetry of 'because' does not rule out response-dependence concepts; it only puts some constraints on them. First, response-dependence as sketched above is no problem as long as the beliefs which make a response-dependent concept apply to something are themselves not due to its applying to the thing (i.e. if a belief that  $x$  is cool, which contributes to making  $x$  cool, is explained by facts other than that  $x$  is cool). Second, asymmetry even leaves room for *some* of the beliefs that make a concept apply (e.g. some of the cool-making beliefs) to be explainable by the fact that the concept applies. Earlier we saw that feedback structures which involve temporal successions of interaction do not violate asymmetry, and such structures can be involved in those cases: someone can be a star because sufficiently many people think her one. On the other hand, people may think her a star because she is one. Their belief that she is a star (now) can

then, without any problems for asymmetry, contribute to making it a fact that she *remains* a star (or: that she is a star at some later time).

The third case is a famous example by van Fraassen (1980, 132–34): someone built a tower wanting it to cast, at a particular time of the day, a shadow reaching some particular spot. The required height of the tower could be calculated from the length that the shadow was intended to cast. Given this story, van Fraassen holds, there are contexts in which the following statement is true:

- (8) The tower is 10 meters high because the shadow is 10 meters long.

Yet, there are clearly also contexts in which the reverse direction of (8) holds:

- (8<sub>R</sub>) The shadow is 10 meters long because the tower is 10 meters high.

Put in the terminology of this paper, van Fraassen suggests that (8) and (8<sub>R</sub>) illustrate that there are two non-harmonious priority relations.<sup>32</sup> However, his example, witty as it may be, remains unconvincing.<sup>33</sup> For in van Fraassen's story, the following explanation of the tower's height is available:

- (9) The tower is 10 meters high because *it was built with the aim of ensuring that* the shadow is 10 meters long.

And while (9) may be unpacked into a more complete explanation by adding bits of relevant information, such as that the tower was built according to the intention, it was an accurate piece of work, etc., what has no role to play in any complete explanation is the information that the shadow actually *is* 10 meters long.

But then we have no reason to think that (8) is literally true with respect to the described context. Regardless, the utterance of (8) may be pragmatically *acceptable* because even though it is actually false, it is easy to compute a true remark—i.e. (9)—out of it. Still, this does not reveal any peculiar semantic feature of 'because'. In conversation,

---

32. For van Fraassen, the example serves additional purposes, such as illustrating the context-dependence of 'because'. Those purposes are irrelevant here.

33. On the following, cp. Horwich (1987, 151f.) and Kitcher & Salmon (1987, 316f.) (who, however, also argue that the above interpretation of van Fraassen, though natural, eventually must be replaced by a different one).

we usually tolerate a considerable amount of slack and imprecision. If someone utters a false sentence, we let it pass if we can easily identify a truth that the speaker may have *meant*. Nevertheless, we *can* correct such imprecision, and a correction of (8) along the lines of (9) should and would be accepted by the speaker (unless he was trying to prove a wrong philosophical point).

Summing up, the examples discussed do not show that the relations potentially grounding because-sentences are disharmonious. Even still, the current worry about the asymmetry thesis is the most vital one. The thesis has to be *qualified* if someone comes up with convincing candidates of non-harmonious grounds of ‘because’, but not given up: In that case, asymmetry has to be restricted to because-sentences based on certain priority relations (i.e. those which actually are harmonious). All told, such a restriction might still be necessary. In the limiting case, if all relevant priority relations were pairwise in disharmony, asymmetry would only prevail as long as a single priority relation is concerned. But there are no indications for such a global disharmony; in fact, we have yet to see any convincing case of disharmony.

#### f. *Correlation and explanation*

Here is another potential threat for the asymmetry of ‘because’:<sup>34</sup> in physical theories, it may happen that we find—at a fundamental level of description—correlated properties *P* and *Q* whose possession co-varies with another. Since the level is fundamental, none of the correlated properties is more basic than the other; none has objective priority over the other. Yet the scenario may seem to license the claim that whatever has the one, has it because it has the other, and vice versa. Hence, if a particle has property *P*, the following sentences are both true:

(10) That particle has property *P* because it has property *Q*.

(10<sub>R</sub>) That particle has property *Q* because it has property *P*.

An argument to this conclusion might be built on the notion of *dependence*: The scenario is a classical case where two values or properties are described as interdependent. But dependence gives rise to explanation, as

---

34. The following is an attempt of interpreting a core aspect of Woodward’s (1984, 433–41) position, though it may very well fail to capture his intentions. But anyway, a similar reasoning has been suggested to me more than once, so it should be discussed here.

is often pointed out in the recent debate about ontological dependence: if  $x$  ontologically depends on  $y$ , then  $x$  exists because of  $y$ . Correspondingly, one might hold, if possession of a property  $P$  depends on possession of  $Q$ , then things have  $P$  because they have  $Q$ .

The problem with this reasoning is that talk about dependence admits of several interpretations. Talk about functional dependence, as it typically occurs in scientific theories, is best understood in terms of co-variance (perhaps, with some sort of modal force, e.g. nomological necessity).<sup>35</sup> But dependence understood as co-variance is a weak notion which does not correspond to any notion of priority; this is the central observation that motivated alternative approaches to notions of dependence, such as essential dependence (Fine 1995), identity-dependence (Lowe 1998), and explanatory dependence (Correia 2005, Schnieder 2006).<sup>36</sup> We need not go into the details of these accounts. Suffice it to say that they construe dependence as the converse of some form of priority; thus statements involving such stronger notions of dependence may indeed imply because-statements. But the scenario described does not establish an interdependence between the two properties in such a strong sense. That we can talk about interdependence here was made plausible precisely in terms of mutual co-variance (correspondence). But co-variational dependence can be granted without granting any because-statement.

A second argument to the effect that the scenario licenses both (10) and (10<sub>R</sub>) might start from the intuition that the correlation of the two properties  $P$  and  $Q$  just seems to have some explanatory value. Hence we can explain possession of one in terms of possession of the other.

Here one may admit that the correlation of the two properties is explanatorily valuable, while holding that this is due to the varying epistemic connotations of 'explain'. Doubtless, noting the correlation between  $P$  and  $Q$  has a great scientific value—a theory which notes it increases our understanding of the involved phenomena and, in some epistemic sense of 'explain', thereby helps in explaining them. But that can be frankly admitted without the acceptance of any non-evidential because-claims ('things have  $P$  because they have  $Q$ '). Of course, in the scenario such claims are acceptable in an evidential reading (note that the evidential reading may in fact be backed by modal interdependence). But those who believe the sentences should be accepted also in a non-evidential sense should provide

---

35. For a seminal discussion of dependence understood as necessary co-variance, see Simons (1987, ch. 8).

36. For a good survey, see Correia (2008).

some reason; after all, the mere constant correlation of two properties does not normally license a because-claim.

Finally, note that a good way of explicating what it means for some properties to be *fundamental* is in terms of 'because'. Roughly, a property is fundamental iff whatever possesses it does not possess it because some other facts obtain. On such an understanding of 'fundamental' it immediately follows that there cannot be two fundamental properties  $x$  and  $y$  such that things have  $x$  because they have  $y$  (or vice versa).

g. *Explanation and Identity*

Another problematic case is suggested by Ruben who thinks *identity explanations* are (the only) genuine violations of the asymmetry of explanation. Here is his idea:

If one assumes as background the theory which identifies various ordinary substances with chemically precise compounds and mixtures, then in the appropriate circumstances, the fact that ice is water can be fully explained by the fact that ice is  $H_2O$ . In other circumstances, the fact that ice is  $H_2O$  can be fully explained by the fact that it is water. It depends on what is known and what needs explanation. (Ruben 1990, 222)

By stressing the importance of what some person knows and what he wants to be explained, Ruben sounds like he is making a point about *explanatory acts*. But he is not: earlier, he distinguishes between a process and a product sense of 'explanation', takes the product sense to be independent of the process sense, and sets the process sense aside.<sup>37</sup>

Unfortunately, Ruben is unclear about what he takes the product sense of 'explanation' to be, and hence what he takes explanations to be. But assume he thinks that a product-explanation is the content of a because-sentence (at least, he often uses 'because'-formulations when he gives examples of explanations). His example is then relevant to the asymmetry of 'because'. Judging from the quotation he would hold that, in certain circumstances, it is true that

(11) Ice is water because ice is  $H_2O$ ,

whereas in other circumstances, it is true that

---

37. Ruben (1990, 6–9).

(11<sub>R</sub>) Ice is H<sub>2</sub>O because ice is water.

The relevant circumstances differ in ‘what is known and what needs explanation’. But how exactly should this feature of the circumstances bear on the truth-values of (11) and (11<sub>R</sub>)? It certainly does so in the evidential sense of ‘because’. But in that sense, the sentences would not be counter-examples to the asymmetry of ‘because’. Moreover, Ruben stresses that identity explanations are peculiar in violating the asymmetry of explanation, but if an evidential ‘because’ was involved, this would just be wrong: as we saw above, the evidential ‘because’ can commonly reverse the non-evidential. But if (11) and (11<sub>R</sub>) are read in a non-evidential sense, then Ruben gives us no reason for thinking that their truth varies with features of any epistemic subjects. Nowhere in his book does he present an account to that effect. So perhaps one must conclude that, in the quotation, Ruben fell into a trap he earlier warned his readers about: getting confused about different senses of ‘explanation’. What he presents may be a good point about explanatory acts (and evidential because-statements), showing that they are not asymmetrical across different contexts. But as a point about the non-evidential ‘because’, it is hard to see how it should work without substantial additional argument.

#### h. Backward Explanations

Our last potential problem case comes from Jenkins and Nolan (for short: J&N) who point out that, in everyday parlance, we sometimes use because-sentences to give *backwards explanations*, which explain the occurrence of an event by the future occurrence of another event. Here is a representative example from their stock:

(12) The scarlet pimpernels closed up because it was about to rain.  
(J&N 2008, 104)

That explanations such as (12) can indeed be true will not be doubted here; in fact, since such cases appear to be widespread one may accept them as relevant data that a semantic theory of ‘because’ must take into account.

J&N contrast their acceptance of backwards explanations with a quoted remark from Strevens on the explanatory asymmetry:

The principal motivation for the causal approach is the recognition [...] that the direction of explanation runs parallel to the direction of causation. The height of a flagpole explains the length of its shadow, but the length of the shadow does not explain the flagpole's height. This explanatory asymmetry mirrors a causal asymmetry. (2004, 156)

It is somewhat surprising that J&N talk of the quotation as containing an explicit commitment to a rejection of backwards explanations. For, the quotation does not mention backward explanation at all, nor does it mention any general temporal patterns of explanation. It simply says that the direction of explanation parallels that of causation, which obviously only applies to cases which involve causation.

Anyway, the quotation mentions an explanatory asymmetry while J&N take the quotation to rule out backwards explanation. The rationale might be that since typical explanations explain events in terms of past events, there being backwards explanations running the other way round should lead to violations of asymmetry.

But do backwards explanations really violate the asymmetry of 'because'? Example (12) certainly does not. For even though many events can be explained in terms of past events, not every past event will do. So, let us consider the reverse explanation of (12):

(12<sub>R</sub>) It was about to rain because the scarlet pimpernels closed up.

This statement may in certain contexts be acceptable, but only in an evidential sense of 'because'. Hence, it is no potential counterexample to the asymmetry of 'because'. Or consider a slightly modified version of (12<sub>R</sub>) which can be used to describe the same state of affairs from a later point in time:

(12<sub>R\*</sub>) It rained (at  $t$ ) because the scarlet pimpernels closed up (at  $t^*$ ).

(Where  $t^*$  is an earlier time than  $t$ .) Again, this is at best acceptable in an evidential sense. None of the everyday examples of backwards explanations given by J&N is such that the corresponding explanation of the later event in terms of the former has any initial plausibility.

But could there be such cases? The question depends on the conditions under which a backwards explanation of the sort 'e occurred because e\*' was going to occur' is acceptable, and settling that issue clearly goes beyond the scope of this paper. Still, a hunch can be given here: one condition under

which a backwards explanation is unacceptable seems to be just when the earlier event *causes* the later event. By way of inductively supporting this hunch, consider some cases: smoking causes cancer, the rain causes the street to get wet, the explosion causes the smoke; ‘he’s smoking because he’s going to get cancer’, ‘it’s raining because the street will get wet’, and ‘the thing explodes because there will be smoke’ are all unacceptable. But it would be exactly such cases—in which a backwards explanation explains a cause in terms of its effect—which were likely to violate the asymmetry of ‘because’. We explain events not by *any* past events, but by those which *caused* them. So, there is at least some reason not to expect any genuine trouble for the asymmetry thesis from backwards explanations of the mundane sort.<sup>38</sup>

### §5. *Concluding remarks*

Let me take stock. As a whole, the paper offers a modest defence of the asymmetry thesis. It started with formulating the thesis and went on to considering whether, if the thesis is true, its truth is a brute fact or whether it can be accounted for in more fundamental terms. Two proposals were found wanting: the account in terms of explanatory acts and the account in terms of the transitivity and irreflexivity of ‘because’. However, an account based

---

38. Apart from their mundane examples of backward explanations, J&N also present an example that ‘appeal[s] to exotic metaphysical possibilities’: the time traveller Dr. Bostrom appears in 1905 in Framley because he built the time machine in 2005. *Per se*, the example again does not yield any plausible candidate for a reversible ‘because’-sentence. But an expanded version of the scenario involving a causal loop may do: Assume that Bostrom only built the machine in 2005 because he found a newspaper article from 1905 describing his own appearance in 1905. Then it may be tempting to accept violations of asymmetry: Bostrom builds the machine in 2005 because he made an appearance in 1905, and he made that appearance in 1905 because he builds the machine in 2005.

Since time travel is a highly controversial and delicate issue, it cannot be treated here (J&N put it aside for the same reason). Let me just point out that even if one takes time travel seriously, the consequences for the asymmetry of ‘because’ would still not be obvious and depend on the exact details of the theory about time travel. To mention one relevant issue: Some philosophers who take time travel to be possible still deny that it can involve causal loops (for references, see Smith 2013); on such a view, time travel would not endanger the asymmetry of ‘because’. It would endanger a claim about the *temporal* profile of ‘because’, namely that if ‘... at *t* because --- at *t*\*’ is true, then *t*\* must be prior to *t*. While that feature might be referred to as the *temporal asymmetry* of ‘because’, it did not play any role in this paper; moreover, a failure of temporal asymmetry is compatible with ‘because’ being an asymmetrical connective in the sense discussed in this paper.

on an Aristotelian semantics of ‘because’ seemed workable (even though it provides an explanans which is quite close to the explanandum). Finally, the paper discussed several potential counterexamples to the asymmetry thesis and developed responses on behalf of the proponent of the thesis.

If the rejections are accepted, where does this leave us? Someone might argue that even if the examples considered can be rebutted, the mere fact that there are so many gives us good inductive evidence not to believe in the asymmetry of ‘because’; convincing counterexamples will certainly be found sooner or later.<sup>39</sup> Two replies: first, there are far too many attractive philosophical theses for which opponents can adduce numerous alleged counterexamples to accept the general methodology suggested by the argument. Second, the reactions to the examples use resources which are available to evaluations of because-statements anyway; and since ‘because’, as a matter of fact, is highly sensitive to minor differences in the formulation of its clauses, easily generates implicatures, has an ambiguity between an evidential and an objective sense, etc. we should *expect* there to be many *prima facie* counterexamples to general claims about its behaviour.

On a concluding note, if there were strong counterexamples against the asymmetry of ‘because’ (properly understood) which cannot be discharged as spurious, why is it that nobody came up with them yet? The proponent of the asymmetry may stick to her thesis until someone invents such examples. Until then, she may even find some reassurance in the fact that, so far, nobody has.<sup>40</sup>

#### LITERATURE

Achinstein, Peter 1983: *The Nature of Explanation*. Oxford: Oxford University Press.

Bender, Emily & Kathol, Andreas 2001: “Constructional Effects of “Just Because ... Doesn’t Mean ...””. *Berkeley Linguistics Society: Proceedings of the Annual Meeting 27*, 13–25.

---

39. The argument was proposed to me by Daniel Nolan.

40. Thanks for discussion and comments to the members of *Phlox*—Miguel Hoeltje, Nathan Wildman, Alex Steinberg, Moritz Schulz, and Nick Haverkamp—, and to audiences in Canberra, Gothenburg, Stockholm, and Lund. Thanks also to Florian Fuchs and Nina Scheller for their help with the manuscript and to the editor of this Festschrift and to her assistant for the work they invested in it.

- Bird, Alexander 2003: "Philosophy of Science". In: John Shand, *Fundamentals of Philosophy*. London: Routledge, 297–325.
- Bolzano, Bernard 1837: *Wissenschaftslehre* (4 vols.). Sulzbach: Seidel. Reprinted in: *Bernard Bolzano Gesamtausgabe* 1.11–1.14. Stuttgart: Frommann-Holzboog (1985ff.).
- Bromberger, Sylvain 1965: "An Approach to Explanation". In: R. J. Butler (ed), *Analytical Philosophy*. Vol. 2, 72–105.
- Brown, John H. 1964: "The Logic of the Euthyphro 10A–11B". *The Philosophical Quarterly* 14(54), 1–14.
- Caputo, Stefano 2007: "Truth-making: what it is not and what it could not be". In: Monnoyer Jean-Maurice (ed.), *Metaphysics and Truthmakers*. Frankfurt a.M.: ontos, 275–311.
- Cohen, S. Marc 1971: "Socrates on the Definition of Piety: *Euthyphro* 10A–11B". *Journal of the History of Philosophy* 9(1), 1–13.
- Correia, Fabrice 2005: *Existential Dependence and Cognate Notions*. Munich: Philosophia.
- 2008: "Ontological Dependence". *Philosophy Compass* 3(5), 1013–1032.
- Correia, Fabrice & Schnieder, Benjamin (eds.) 2012a: *Metaphysical Grounding*. Cambridge: Cambridge University Press.
- 2012b: "Grounding: An Opinionated Introduction". In: Fabrice Correia & Benjamin Schnieder, *Metaphysical Grounding*. Cambridge: Cambridge University Press, 1–36.
- Fine, Kit 1995: "Ontological Dependence". *Proceedings of the Aristotelian Society* 95, 269–90.
- 2012: "Guide to Ground". In: Fabrice Correia & Benjamin Schnieder, *Metaphysical Grounding*. Cambridge: Cambridge University Press, 37–80.
- Frankel, Louis 1986: "Mutual Causation, Simultaneity and Event Description". *Philosophical Studies* 49(3), 361–72.
- Gärdenfohrs, Peter 1980: "A Pragmatic Approach to Explanations". *Philosophy of Science* 47(3), 404–23.
- 1990: "An Epistemic Analysis of Explanations and Causal Beliefs". *Topoi* 9(2), 109–24.
- Grice, H. P. 1967/68: "Indicative Conditionals". In: H. P. Grice 1989: *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press, 58–85.
- Hall, John C. 1968: "Plato: Euthyphro 10a1–11a10". *The Philosophical Quarterly* 18(70), 1–11.
- Hausman, Daniel H. 1982: "Causal and Explanatory Asymmetry". *PSA* 1982, vol. I, 43–54.
- 1993: "Linking Causal and Explanatory Asymmetry". *Philosophy of Science* 60(3), 435–51.

- Hesslow, Germund 1981: "The Transitivity of Causation". *Analysis* 41(3), 130–33.
- Hornsby, Jennifer 2005: "Truth Without Truthmaking Entities". In: Helen Beebe & Julian Dodd (eds.), *Truthmakers—The Contemporary Debate*. Oxford: Clarendon Press, 33–47.
- Horwich, Paul 1987: *Asymmetries in Time*. Cambridge, Mass.: MIT Press.
- Jenkins, Carrie & Nolan, Daniel 2008: "Backwards Explanation". *Philosophical Studies* 140(1), 103–15.
- Kac, Michael 1972: "Clauses of Saying and the Interpretation of *because*". *Language* 48(3), 626–32.
- Kenny, Anthony J. P. 1998: *A Brief History of Western Philosophy*. Oxford: Blackwell.
- Kitcher, Philip & Salmon, Wesley 1987: "Van Fraassen on Explanation". *Journal of Philosophy* 84(6), 315–30.
- Künne, Wolfgang 2003: *Conceptions of Truth*. Oxford: Clarendon Press.
- Lange, Marc 2009: "Why Proofs by Mathematical Induction Are Generally not Explanatory". *Analysis* 69(2), 203–11.
- Liggins, David 2008: "Truthmakers and the Groundedness of Truth". *Proceedings of the Aristotelian Society* 108(2), 177–196.
- Litland, Jon 2013: "On Some Counterexamples to the Transitivity of Grounding". *Essays in Philosophy* 14(1), 19–32.
- Lowe, E. J. 1998: *The Possibility of Metaphysics*. Oxford: Clarendon Press.
- Mohamed, Mohamed Taha & Clifton, Charles 2008: "Processing Inferential Causal Statements: Theoretical Refinements and the Role of Verb Type". *Discourse Processes* 45(1), 24–51.
- Morreal, John 1979: "The Evidential Use of *Because*". *Papers in Linguistics* 12(1–2), 231–38.
- Mulligan, Kevin 2006: "Ascent, Propositions, and Other Formal Objects". *Grazer Philosophische Studien* 72, 29–48.
- Newton-Smith, W.H. 2000: "Explanation". In: W.H. Newton-Smith: *A Companion to the Philosophy of Science*. Oxford: Blackwell, 127–33.
- Nozick, Robert 1981: *Philosophical Explanations*. Cambridge, Mass.: Harvard University Press.
- Paxson, Thomas D. 1972: "Plato's Euthyphro 10a to 11b". *Phronesis* 17(2), 171–90.
- Putnam, Hilary 1973: "Reductionism and the Nature of Psychology". *Cognition* 2(1), 131–46.
- Railton, Peter 1980: *Explaining Explanation: A Realist Account of Scientific Explanation*. Unpublished Ph.D. Dissertation, Princeton University.
- Rodriguez-Pereyra, Gonzalo 2005: "Why truthmakers". In: Helen Beebe & Julian Dodd (eds), *Truthmakers: The Contemporary Debate*. Oxford: Clarendon Press, 17–31.

- Rosen, Gideon 2010: "Metaphysical Dependence: Grounding and Reduction". In: Bob Hale & Aviv Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford: Oxford University Press, 109–36.
- Ruben, David-Hillel 1990: *Explaining Explanation*. London: Routledge.
- Schaffer, Jonathan 2012: "Grounding, Transitivity, and Contrastivity". In: Fabrice Correia & Benjamin Schnieder 2012a, 122–38.
- Schnieder, Benjamin 2004: "Compatibilism and the Notion of Rendering Something False". *Philosophical Studies* 117(3), 409–428.
- 2006: "A Certain Kind of Trinity: Dependence, Substance, Explanation". *Philosophical Studies* 129(2), 393–419.
- 2010: "A Puzzle About "Because"". *Logique et Analyse* 5, 317–43.
- 2011: "A Logic for "Because"". *Review of Symbolic Logic* 4(3), 445–65.
- Schnieder, Benjamin & Steinberg, Alex (forthcoming): "Without Reason?". *Pacific Philosophical Quarterly*.
- Searle, John 1969: *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Sharvy, Richard 1972: "Euthyphro 9d-11b: Analysis and Definition in Plato and Others". *Noûs* 6(2), 119–37.
- Simons, Peter 1987: *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- Smith, Nicholas J.J. 2013: "Time Travel". *The Stanford Encyclopedia of Philosophy* (Winter 2013). Edward E. Zalta (ed.), URL =<<http://plato.stanford.edu/archives/win2013/entries/timte-travel/>>
- Strawson, Peter F. 1957: "Propositions, Concepts, and Logical Truths". *Philosophical Quarterly* 7(26), 15–25.
- Strevens, Michael 2004: "The Causal and Unification Accounts of Explanation Unified—Causally". *Noûs* 38(1), 154–179.
- Tatzel, Armin 2002: "Bolzano's Theory of Ground and Consequence". *Notre Dame Journal of Formal Logic* 43(1), 1–26.
- Trout, J. D. 2002: "Scientific Explanation and the Sense of Understanding". *Philosophy of Science* 69(2), 212–33.
- van Fraassen, Bas 1980: *The Scientific Image*. Oxford: Oxford University Press.
- Williamson, Timothy 1999: "Truthmakers and the Converse Barcan Formula". *Dialectica* 53, 253–270.
- Woodward, James 1984: "Explanatory Asymmetries". *Philosophy of Science* 51(3), 421–42.

NON-TRANSITIVE PARTHOOD, LEVELED MEREOLGY,  
AND THE REPRESENTATION OF EMERGENT PARTS  
OF PROCESSES

Johanna SEIBT  
Section for Philosophy and the History of Ideas  
Aarhus University

*Summary*

Processes have mereological structure, just like things and stuffs, but part-whole relations for processes have not received much attention in the literature so far. As I explain in the first part of this paper, the main reason for this curious neglect is that extant classical and non-classical mereologies have hidden built-in restrictions on the type of entities that can stand in the part-relations formalized by these mereologies; processes and other non-particular individuals do not fulfill the given restrictions. In a second step I introduce a non-classical mereological system (LEM) which is free of these restrictions and operates with a nontransitive 'is-part' relation, capturing the most general sense of mereological association: 'belongs-with.' In a third step I discuss to what extent LEM can be used to formally represent our qualitative reasoning about processes.

Processes have mereological structure, just like things. Unlike things, however, processes have parts that can be said to 'emerge,' in a fairly strong sense of that term. Curiously, however, the debate about emergence, both the older debate in the 1950s and the more recent discussion since 1990, has largely been conducted with focus on emergent *properties* of things rather than on emergent (parts of) processes. In fact, part-whole relations for processes altogether have not received much attention in the literature so far.<sup>1</sup> Why have we all but forgotten about processes in analytical ontol-

---

1. As in many other areas in ontology, Peter Simons' short discussion of the topic in his 1987 (129-147) offered a seminal pointer by creating an association between the formal mereology of stuffs and the mereology of processes, which, in informal terms and with focus on the semantics of verbal aspects, had been observed by Dowty (1977) and, in particular, Mourelatos (1978) .

ogy, and in the debate about emergence and in mereology in particular? What would it take to devise a theory of parts for processes, and which theoretical benefits might we gain from it, besides extending the scope of ontological analysis? These questions lead into fairly unexplored terrain and the following considerations will do little more than point to some possible trailheads towards a non-Whiteheadian process ontology. If the following sketch will not succeed in recruiting fellow explorers, if it merely will serve to redirect readers towards the highway that Whitehead blasted into the jungle of metaphysics and hid under the overgrowth of foreign terminology, they will have fulfilled a dialectical function that is not entirely unwelcome.<sup>2</sup>

In the first part of this paper I question the ubiquitous belief that ‘part-of’ in its most basic sense is a transitive relation. I suggest that there is a joint reason for the neglect of processes in mereology and for the dogma of transitivity—these are, among many others, the effects of a longstanding theoretical tradition in ontology that even the ‘analytical’ reformation of the discipline in the first half of the 20<sup>th</sup> century could not shed. The theoretical presuppositions of this tradition—which I call the ‘substance paradigm’ or the ‘myth of substance’—generate restrictions on the types of entities that can count as ontologically basic, as well as on the types of entities that are primary targets of ontological investigation. In consequence of this restriction on entity types, a certain type of part-whole relation took center-stage in the development of mereology: spatial parthood on spatial regions. The transitivity of this sort of parthood also abetted the idea that mereology and emergence are anathema, and that one cannot model emergence using mereological relationships. In a second step I sketch a non-classical mereological system (LEM: Leveled Mereology), which operates with a non-transitive ‘is-part’ relation and is designed to capture the logical properties of the most general and inferentially weakest part-whole relation: ‘belongs-with.’ In a third step I show how claims about the emergent parts of processes can be formulated with the tools of LEM. I conclude with a brief discussion of the significance of this formal presentation.

---

As far as I can see, I may have been the first one (cf. Seibt 1990, ch. 5) to pursue this pointer systematically and to further explore the mereology of processes alias “dynamic masses”: cf. the references listed on pages 174 and 183 below.

2. C.f. e.g., Simons (2000).

## 1. *Classical transitive parthood and the “myth of substance”*

The processing of part-whole relationships seems to be one of our most basic cognitive capacities. In terms of the partly phenomenological, partly logical analysis of cognition that philosophers used to call ‘transcendental’ and still often undertake, it would seem possible to argue, for example, that part-whole relationships are cognitively more fundamental than member-set relationships. The cognitive ‘grasp’ of member-set relationships, one might say, presupposes the capacity to focus on some feature of an entity and to ‘represent’ the entity in terms of that feature and to associate it with others selectively ‘represented’ by similar features.<sup>3</sup> Traditionally, this cognitive capacity of selective representation has been called ‘abstraction,’ retaining the terminology, if not the theoretical model, of the Scholastic account of perception. But one might also argue that the selective representation that is prerequisite for member-set relations is connected to the linguistic phenomenon of *metonymy* or, more precisely, *synekdoche*. In *synekdoche* an entity E is referred to by using an expression L which denotes a part of that entity, e.g., ‘I need a new *set of wheels*,’ or: ‘what we want are more *bums* on seats’; alternatively, L denotes the whole of which E is a part, e.g., ‘Do you need to use *the bathroom*?’ or ‘I am going to fill up *the car* with petrol.’<sup>4</sup> *Synekdoche*, one might suggest, manifests at the linguistic level an underlying cognitive ability to grasp ‘at a glance’ a situation with different features and a basic relationship of asymmetric ‘belonging with,’ so that the simple or complex item on either side of this relationship can be foregrounded to refer to the other. In short, guided by the linguistic phenomenon of *synekdoche* one might hypothesize that member-set relationships are cognitively grounded in referential ‘mechanisms’ where features of a situation come to stand for the whole situation, or vice versa.

Whether the hypothesis that part-whole relationships are cognitively prior to member-set relationships can be argued for along these lines would need to be investigated in cognitive psychology or cognitive science. I have introduced it here since it draws attention to a feature of part-whole relationships that can be easily confirmed by introspection. It is phenomenologically striking and systematically important, in my view, that part-whole relations, *some* at least, can be grasped ‘at a glance,’ immediately—we seem to be able to directly perceive them or at least process them without

---

3. In the context of this paper I want to bracket the question which, if any, notion of mental representation should still be used in present-day philosophy.

4. Cf. Rijkhoff (2015), whence also all sample sentences are taken.

any noticeable inferential step. This phenomenologically non-inferential character of the part-whole relation holds in particular for *spatial parts*. Whether something is spatially part of something else we simply *see*—we take it in as an *integral* aspect of the scenery presented.

Let (A1) be the assumption that phenomenal immediacy (or non-inferentiality) is a criterion (i.e., necessary condition) for the cognitive basicness of a relationship. The above argument for the cognitive priority of parthood hinges on (A1) and the phenomenological datum of the phenomenal immediacy of spatial parthood. In contrast, membership requires more complex processing and is not phenomenally immediate; moreover, membership seems to involve the processing of spatial parthood in one of its constitutive components, that is, the referential shift manifested in *synekdoche*, from whole entities to features. Encouraged by this line of reasoning one might further be tempted to argue that foundational theories in analytical ontology should not resort to set-theory but use classical extensional mereology as a basic formal framework. More explicitly, one might try to establish such a recommendation on the basis of (A1), the immediacy criterion for cognitive basicness, and the following additional assumptions:

- (A2) Foundational ontologies should aim to operate with formal frameworks that represent cognitively basic relationships.
- (A3) Spatial parthood is a cognitively basic relationship.
- (A4) Classical extensional mereology (CEM) models spatial parthood *per se*, i.e., spatial parthood on arbitrary types of entities.
- (A5) (CEM) models the most generic notion of parthood that is cognitively basic in the sense of (A1), i.e., there is no other cognitively basic part-whole relationship that is more generic or comprehensive than spatial parthood *per se* and is not modelled by (CEM).

I suspect that assumption (A4) will appear straightforward to many mereologists.<sup>5</sup> To be sure, due to Simons' pioneering 1987 there is general acknowledgement among mereologists that (CEM) cannot model all part-relations, especially those that carry modal information, such as constitutive parthood or material constitution. But modal part-whole relationships are precisely not phenomenologically immediate, and so one might hold on to the idea that (CEM) captures at least spatial parthood, the cognitively basic parthood relation. Significantly, the basic principles

---

5. For a clear statement of this commitment cf. e.g., Varzi 2006:1.

of (CEM) are typically illustrated with geometric diagrams of spatial inclusion. What has gone unnoticed, however, is that it is not spatial parthood *per se* that provides a model for (CEM) but spatial parthood *on spatial regions*. Objects, stuffs, activities, events, persons, collectives etc. all have spatial parts, but these do not coincide with the items occupying regions that are spatial parts of the spatial region occupied by the object. Consider the following sentences, letting ‘S’ denote a certain screw that fixes a cable running inside one of the walls of my house:

- [1] A screw is a spatial part of a house. (false)
- [2] S is a spatial part of my house. (?)
- [3] The spatial region occupied by S is a spatial part of the region occupied by this house. (true)

Sentence [1], which is a generic statement, is false since screws are not typically used in the construction of houses. The falsity of [1] however disinclines us to treat [2] as true, especially if we are also provided with sentence [3] as a clarification of why we initially might have considered [2] as true (what we would have meant had we accepted [2] as true).<sup>6</sup> But if [2] is false, spatial parthood on objects such as houses is not a strictly transitive relation as required by (CEM)—the wall is a spatial part of my house but something spatially in the wall apparently is not.

Similar counterexamples to the transitivity of spatial parthood can be constructed for other kinds of objects:

- [4] An influenza virus is a spatial part of a human organism. (false)
- [5] (I have influenza and carry virus V.) V is a spatial part of my body. (?)
- [6] The spatial region occupied by V is a spatial part of the region occupied by my body. (true)

The example is to show that something spatially included in the spatial part of a thing may fail to qualify as a spatial part of that thing. As illustrated

---

6. This is an empirical claim, as is any claim about ‘what we would say’ or ‘would not say’, i.e., about our normative dispositions. Note that temporal relativizations of the second and third sentences (“At t, ...”) are omitted, here and in the examples below. This and the following examples can also serve to make a methodological point in passing. I do not think that our so-called “metaphysical intuitions” are stable across contextualizations—as sociologists and psychologists are well aware of, already the order in which certain information is presented affects a subjects responses.

by sentences [1] through [6], the ‘transitivity failure’ of spatial parthood on objects arises due to the fact that our nouns for objects are abbreviations for functional organizations. Objects such as houses, desks, doors, or dishwashers have functional organizations (functional structures), which reflect the dispositional features of objects relative to physical requirements and/or socio-historical contingencies. Whether something is a spatial part of an object we determine relative to the functional organization of this kind of object. In general, if we determine the spatial parts of an object O, spatial parthood is transitive as long as it pertains to items that belong to the functional organization of O.<sup>7</sup>

Problems with the transitivity of spatial parthood have been discussed since Aristotle but mostly with focus on the alleged spatial “granularity” of stuffs (‘is one oxygen molecule a spatial part of water?’) and the alleged temporal granularity of activities (‘is my inhaling part of my singing?’).<sup>8</sup> It is important, however, to separate the discussion of the granularity of stuffs and processes from transitivity failures for spatial parthood on stuffs and processes that arise for claims about spatial items that are not included in the functional organization of the stuff or process in question. Consider again the following triplets of sentences:

- [7] Radon is a spatial part of gin-and-tonic. (false)
- [8] (The tonic water I used for my drink is contaminated with certain amount of radon, call it R.) R is a spatial part of my (glass of) gin-and-tonic.
- [9] The (scattered) spatial region occupied by R is a spatial part of the region occupied by my (glass of) gin-and-tonic. (true)
- [10] The pistons’ rattling is a spatial part of an engine’s running. (false)
- [11] My car is very old and the pistons in my car’s engine rattle; call that activity of rattling ‘R’ and the activity of my (car’s) engine’s running ‘E’. R is a spatial part of E’s running.
- [12] The spatial region in which the rattling and knocking occurs is a spatial part of the spatial region in which E occurs. (true)

---

7. Matters are complicated by the fact that the terms for objects are often connected with several default organizations—for example, the ridge beam is a spatial part of a house, even though many houses do not have ridge beams.

8. While so-called ‘granularity arguments’ serve well to highlight transitivity failures of spatial parthood on arbitrary types of entities besides spatial regions, it is important to note that they involve a clandestine “aspect shift” from stuffs to mixtures and from activities to sequences of events, cf. Seibt (2004b, ch. 4).

Sentences [8] and [11] will be judged false, I submit, once contrasted with [9] and [12], the only reading in which they could have appeared to be true—a certain amount of water *W* is a spatial part of the entity referred to by ‘the gin-and-tonic in my glass’ (whose functional structure *F* determines what counts as spatial part), but the amount of radon contaminating the water is at best a spatial part of the spatial region occupied by the mixture of chemical substances in my glass; similarly, the rattling *R* is at best a spatial part of the spatial region occupied by the complex of processes occurrent under the hood of my car, but not a spatial part of the running of my car’s engine.

These transitivity failures of spatial parthood on objects, stuffs, and activities show that the strictly transitive part-relation of (CEM) cannot count as a formal reconstruction of spatial parthood *per se*, i.e., irrespective of the ‘relata’ of spatial parthood; at best, (CEM) formally captures *spatial parthood defined on spatial regions*. So we should reject (A4).

If (CEM) is not a formal reconstruction of spatial parthood *per se*, however, it also follows that assumption (A5) should be rejected: the part relation formally defined by (CEM) does not cover any part relation that is cognitively basic. But besides the illustrated transitivity failures of spatial parthood on objects, stuffs, and activities, there are also other reasons for giving up on the idea that a cognitively basic parthood relation is strictly transitive, as (CEM) requires. For perceptual immediacy in the sense of non-inferentiality holds not only for part-relations that we visually perceive, but also for part-whole relationships that we hear, smell, feel by touch, or are proprioceptively aware of. I hear the flute *in* the sound of the orchestra’s playing; I smell ‘apple’ *in* the fragrance of your perfume; I feel the smoothness and the softness *in* the texture of my cat’s fur; I feel how I move each leg *in* my walking. As mentioned above, phenomenologically speaking there seems to be a basic relationship of asymmetric ‘belonging with’ that holds across sensory modalities. But such basic, directly perceived ‘belonging with’ is subject to the context-sensitivity of perception and thus is not strictly transitive. When the flutist plays alone, the sound I experience has overtones, i.e., the overtones belong with the sounds she makes; but when the flutist plays the same melody within the orchestra, these overtones do not belong with the sound I experience from the orchestra. Similarly, as every cook knows, the taste of chili, which belongs with hot pepper eaten in isolation, disappears within curry with much cream and pineapple. I feel my weight shifting if I lift my left leg, but the weight shift does not belong with my proprioceptive experience of walking. Some

of these transitivity failures are due to contextivity or threshold effects in our sensory experience, some are due to physical or chemical interactions; for present purposes it only matters that the asymmetric ‘belonging with’ relationship that structures what we immediately experience in all sensory modalities is not strictly transitive.

In sum then, if one were to assume (A1) and link cognitive basicness to perceptual immediacy, we could indeed argue that there is a cognitively basic relationship of asymmetric ‘belonging with,’ but since this relationship is not transitive, we could not claim that (CEM) is a formal theory of such a generic basic part-whole relationship. At best, (CEM) is a formal theory of the application of a specific variety of this relationship: spatial parthood, to a specific domain: spatial regions, and the transitivity of spatial parthood on spatial regions results from the identity conditions of spatial regions, which are particularly weak.

As indicated by the subjunctive brackets of the considerations presented so far, I want to leave open here whether (A1) is a reasonable path to cognitive basicness, and, in fact, whether cognitive basicness is a useful criterion for the choice of the foundational relation in an ontological theory. My main reason for introducing this line of argument was to shake up the ubiquitous conviction among ontologists that the basic sense of ‘part’ is transitive and that (CEM)—while not sufficient to reconstruct all sorts of part-whole relationships—formally captures the basic sense of part (Varzi 2006:2). In his classic *Parts* Simons (1987) states that anyone who “seriously disagrees” with the formal properties of the part-relation: irreflexivity, asymmetry, and transitivity, “has failed to understand the word” (11) but at the same time stresses that “‘part,’ like other formal concepts, is not univocal, but has analogous meanings according to whether we talk of individuals, classes, or masses” (2). The general theoretical landscape he draws up is that there is a basic transitive part relation which in application to some domains (“classes and masses”) fulfills an “extensionality axiom” (the Proper Parts Principle (PPP): that the item ‘A’ refers to has the same parts as the item referred to by ‘B’ implies that ‘A’ and ‘B’ refer to the same entity, and *vice versa*); in application to other domains, such as continuants and integral wholes, this basic transitive part-relation is not “extensional” in the sense of the Proper Parts Principle. In this way any variations in the meaning of ‘part’ are associated with the identity axiom PPP and not with the transitivity axiom, which he takes to be indispensable.

There is an alternative, however. If the task of mereology is to articulate the logic of ‘is part of’ as reflected in common-sense and scientific

reasoning, and if we find that the reasoning samples we work from imply that ‘two’ entities (e.g., a person and her body) have the ‘same’ parts or that ‘one’ entity (at different times) has different parts, we can attribute these phenomena either to the fact that parthood is not extensional in this domain *or* to the fact that it is not transitive. Consider the following example. One might argue that PPP does not hold since a person and her body are different entities but have the same parts. Presenting this example, which is due to F. C. Doepke, Simons explains:

The difference between a person and body is that the former only exists as long as certain kinds of process are going on in the body. These processes are not parts, nor are they relationships among the parts, though they are closely connected with such relationships (114).

As Simons’ explanation brings out very clearly, counterexamples against PPP depend on which items we consider to be the parts of the items involved. However, to get the given counterexample to work the (i) clandestine restriction on the types of parts (e.g.: no processes) is actually not enough; one also needs to (ii) equate ‘part’ with ‘spatial part’ *and* (iii) *equivocate* ‘person’ with ‘spatial region occupied by a person,’ as well as ‘human body’ with ‘spatial region occupied by a human body,’ respectively. If the conceptual slide in (iii) were not present, the example would cease to be counterexample to PPP, since a person has as spatial parts all those parts of her body she can have concerns about (feel pain in or feel shame or pride about) but not the millions of cells that are part of her body.

In general, I think that all currently discussed putative counterexamples against the PPP make use of the transitivity of the part-relation and depend on two conceptual manipulations: (i) a reading of ‘is part of’ as ‘is spatial part of’ and (ii) an equivocation of the items involved with their spatial extensions.<sup>9</sup> If we refrain from these conceptual manipulations and drop the idea that ‘is part of’ is a transitive relation, we can retain the PPP.<sup>10</sup> Differently put, the inferential restrictions that we achieve by dropping PPP

---

9. Cf. Seibt 201+a.

10. Similarly for alleged counterexamples to PPP showing that it does not follow that A and B are different if A and B do not have all parts in common—e.g., the notorious hapless cat that loses its tail in an accident. To receive a counterexample here one needs to (i) again equate part with ‘spatial part’ and (ii) argue that A (the ‘cat-at-t’) has spatial-part *tail*, while B (the ‘cat-at-t\*’) does not, and that A and B count as one and the same cat since they are parts of the life of that cat (according to the identity principle for continuants, a continuant is identical with all parts of its life—the cat is “wholly present” (D. Lewis) at any moment of its life). Again, if we reject that a spatial part of a part of the cat’s life is a spatial part of the cat, the counterexample collapses.

as a principle of identity can also be introduced if we drop the transitivity of parthood while retaining PPP. Before illustrating this latter option in the next section, let me conclude this section with a brief reflection on why most ontologists wish to hang on to the former option, i.e., endorse the transitivity of parthood and sacrifice the PPP.

As suggested above, if cognitive basicness is linked to phenomenal immediacy or non-inferentiality, there are part-whole relationships that have just as good a claim to cognitive basicness as spatial parthood on spatial regions, i.e., the relation that serves as prime illustration for the axioms of CEM. Spatial parthood on spatial regions is transitive, other phenomenally immediate part-relations—e.g., spatial parthood on things—are not transitive. Thus the focus on transitive parthood cannot be explained in terms of phenomenal immediacy alone. But why, then, has spatial parthood *on spatial regions* taken centerstage in mereology? Why is it, that even where considerations of phenomenal immediacy and cognitive basicness do not come into play at all, mereologists of all stripes—those who drop PPP, or those who acknowledge that in order to model material or functional parthood transitivity restrictions must be introduced—insists that the “basic” part-relation is transitive?<sup>11</sup> Why is it that in the debate about transitivity of parthood, to the extent to which there has been any debate at all, proponents of transitivity charge opponents of transitivity with equivocating ‘part’ with ‘functional part,’ while the equivocations of the proponents of transitivity I have been highlighting here, of ‘part’ with ‘spatial part’, and of ‘N’ with ‘spatial region of N,’ have not yet been observed?<sup>12</sup> Or again, why is it that the shortcomings of (CEM) for the modeling of the many different kinds of part-whole relationships—such as material parts, functional parts, construction parts, maintenance parts, design parts<sup>13</sup>—have been almost exclusively addressed by relinquishing PPP, without explicit discussion that this strategy is just *one* of the options? Why has the theoretical juncture: ‘extensionality’ (or better: identity via parthood) versus transitivity’ never come into sight?<sup>14</sup>

---

11. See e.g. Guizzardi (2009).

12. For the debate about transitivity see e.g. Cruse (1979), Winston (1987), Seibt (1990, ch. 5; 2004, ch. 4), Varzi (2006), Vieu (2006), Guizzardi (2009); for a further discussion of the equivocations involved in the transitivity debate see Seibt forthcoming.

13. On the varieties of parthood see Cruse (1979), Winston et al. (1987), Simons/Dement (1996), and Simons (2013).

14. I may have been the only one so far who has explored this route, since 1990 in a continuous series of papers devoted to exploring the mereology of processes (see e.g., 1995, 1996, 1999, 2000, 2001, 2004a, 2008, 2009, and in particular 2004b).

A proper answer to these questions—which I can here only point to—would have two parts. The first part would reconstruct the historical origins of mereology during the first decades of the 20<sup>th</sup> century, when mathematicians, metaphysicians, and ontologists (S. Lesniewski, A.N. Whitehead, N. Goodman, S. Leonard and others) searched for a formal theory that could replace set-theory in mathematics (geometry), logic, and the formal reconstructions of conceptual content in the sciences. Since the subset relation is transitive, a transitive parthood relation seemed most suitable; in addition, given the new program of rendering philosophical discourse more formal, basic domain relations needed some inferential strength. But the history of mereology can only explain the motivations for preferring a transitive relation. In order to explain why ‘spatial part of X’ could be equivocated with ‘spatial part of the spatial region occupied by X’ we would need to reconstruct, second, a longstanding research paradigm of the ontological tradition. This paradigm has its roots in the Aristotelian focus on substances, concrete particular individual things that persist in time, but received its decisive impulse in the Cartesian equation of material substances, *res extensa*, with geometric regions. If the spatial extension of an individual object is identified with a continuous geometric region, spatial parthood on objects clearly would appear to be transitive. But, as our examples above should have brought out, while an object extends over a spatial region, not every item in that region (i.e., not every item that is a spatial part of that region) is a spatial part of the object.

In short, in connection with the longstanding theoretical fixation of (Western) metaphysicians on enduring objects or substances, the Cartesian notion of geometric regions as the extensions of objects became an integral component of a comprehensive mindset that influenced the development of formal theories in philosophy, and to the present day dominates in the discussion of analytical ontology. The dogma of the transitivity of spatial parthood *per se* is just *one* reflection of the workings of the “myth of substance”—the presupposed ontological priority of objects—in combination with the Cartesian geometricalization of extensions, which also played into the hands of the idea that individual objects are ‘particulars,’ i.e., that they are *necessarily* uniquely located entities whose *thisness* (identity, individuality) is defined in terms of their location (at a time).<sup>15</sup> If the identity

---

15. The “myth of substance” or the “substance paradigm” can be reconstructed as a network of about 20 constraints on category construction (see my 1990, 2005, 2008, and 2010). Lorenz Puntel identified the “object-ontological dogma” in analytical ontology already in the early 1980s;

of an entity is given by its location, it can indeed seem straightforward to identify the extension of an entity with the geometric region occupied by that entity.

This move loses its plausibility once we turn away from substances and consider activities, such as *the snowing in Aarhus on February 2<sup>nd</sup>, 2014* or *the Royal Orchestra's playing on the Queen's 70<sup>th</sup> birthday*, i.e., entities that are spatially uniquely located but whose identity is not defined in terms of their location (for instance, because they are spatially superposable). These activities occupy (at each time  $t$  in temporal interval  $T$ ) a geometric region  $R$ , where  $R$  can be understood either as the collection of scattered regions that each are continuous and simply connected (the collection of regions occupied by each snow flake or each musician); alternatively,  $R$  can be understood as one multiply connected continuous region (the entire region with spatial holes where at  $t$  no snow flake or musician is positioned). But no matter which of these two interpretations of  $R$  we adopt, since we determine the spatial parts of a snowing or of an orchestra's playing in relation to functional organization of these activities, clearly not every spatial part of the region  $R$  is a spatial part of these activities. The space occupied by a quarter of the nucleus of an H-atom in the water molecule of a snowflake is not a spatial part of the snowing (though surely a spatial part of the spatial region occupied by the snowing).

In sum, then, if mereology is the philosophical analysis of part-whole relationships by means of a formal theory, then mereology has always been characterized by a basic theoretical juncture. The first option is to operate with a transitive part-relation and abandon the identity principle PPP to block unwanted identifications, rejecting what we have come to call the "extensionality" of classical mereology. The second option is to retain PPP but to block "extensionality" much earlier, by abandoning three ideas: (i) the idea that the extension of an item is the geometric region occupied by the item, (ii) the idea that the only basic sense of parthood is 'spatial part' since it is cognitively basic in some sense, and (iii) the idea that spatial parthood *per se*, on arbitrary domains, is a transitive relation. The myth of substance has obscured the second option; as I will now argue, it is an option worth exploring.

---

for the first systematically worked out proposal of a metaphysics free from the relevant dogmatic restrictions see Puntel (2010).

## 2. *Leveled mereology*

As suggested above, there is a phenomenally immediate and, in this sense at least, cognitively basic relationship of asymmetric ‘belonging with’ that structures what we perceptually take in. Such asymmetric ‘belonging with’ is what we express, I want to suggest, when we use ‘part of’ in everyday speech, as a relationship that can be applied to any entity type: things, institutions, activities, events, actions, stuffs, collections, features.<sup>16</sup> ‘Belonging with’ is the most generic and poorest sense of ‘part of,’ an asymmetric association of items in the widest sense that can be further specified in spatial, temporal, material, functional, operational regards. Formally this basic and generic sense of ‘is part of’ as ‘belongs with’ must be formally modeled by a non-transitive relation—there are some transitivities along the chains of ‘belongs with’ but the depth of parthood chains where the inferential pattern of transitivity can be applied depends on the specification of the part-relation involved (spatial, temporal, material, functional etc.) and on the kind of *relata*.

In tandem with the theoretical juncture mentioned above, i.e., the choice between ‘retain transitivity and drop PPP’ versus ‘retain PPP and drop transitivity,’ there is a choice between two strategies for the modeling of transitivity restrictions for specifications of the part-relation, e.g., ‘functional part’ or ‘construction part.’ We can—and this is the common strategy among mereologists currently—begin with a transitive part-relation and restrict its transitivity for certain specifications and domains.<sup>17</sup> Alternatively, we can operate with a non-transitive part-relation and introduce transitivity for certain specifications and domains (in fact, in my view, only for the very limited case of spatial parthood on spatial regions).

The formal framework of ‘belongs with,’ or of ‘part of’ in its most basic and generic sense, operates best, in my view, with the second strategy. Let us look at some basic axioms and definitions of such a frame-

---

16. Compare the following examples, taken from the web: “Blogging is part of life”; “Russia is part of the West”; “Music is part of God’s Universe”; “All I see is part of me” (book title); “Learning to negotiate is part of the advocacy process”; “My heritage is part of who I am”; “Is part of your college education missing?”; “Having a good signed contract is part of running a professional business”; “Boston is part of the Islam”; “Syria is part of the solution”; “Looking immaculate is part of what I do”; “Pain is part of running a marathon”; “Hopping too is part of running”; “Fab Face is part of Screaming Talent”; “The concert is part of the 11th Ludwig v. Beethoven Easter festival.”

17. Cf. e.g., Simons (1987; 2013); Varzi (2006).

work, which I call ‘Leveled Mereology’ (LEM).<sup>18</sup> The part-relation of LEM, symbolized by ‘ $\hookrightarrow$ ,’ is non-transitive (Ax1), asymmetric (Ax2), and irreflexive (Ax3):

$$\begin{aligned} \text{(Ax1)} & \neg[(x \hookrightarrow y \ \& \ y \hookrightarrow z) \rightarrow x \hookrightarrow z] \\ \text{(Ax2)} & (x \hookrightarrow y) \rightarrow \neg(y \hookrightarrow x) \\ \text{(Ax3)} & \neg(x \hookrightarrow x) \end{aligned}$$

Since ‘ $\hookrightarrow$ ’ is non-transitive according to (Ax1), the relationships established by (Ax2) and (Ax3) reach only the ‘immediate parts’ of an entity, that is, the parts at the first level of the partition of an entity. For example, a whole  $\alpha$  cannot be part of itself, in accordance with (A3), but it may well be part of ... part of itself. Similarly, if  $\alpha$  differs from  $\beta$  and has  $\beta$  as part,  $\alpha$  cannot be part of  $\beta$ , in accordance with (A2), but  $\alpha$  may well be part of ... part of  $\beta$ . As I shall briefly elaborate below, the particular interest of LEM lies precisely in the fact that it allows for such ‘parthood loops,’ but let us first consider some definitions that will allow us to extend the logical properties of ‘ $\hookrightarrow$ ’ in a differentiated fashion. Every basic term (individual constant) of (LEM) is associated with a default partition  $P(x)$ ; e.g., let the term ‘ $\alpha$ ’ be associated with the partition  $P(\alpha)$ , which has three levels<sup>19</sup>:

---

18. I have envisioned and sketched Leveled Mereology for some time, always in application to processes (cf. Seibt 1990 ch. 5; 2001, 2004 [with unfortunate misprints in the formulas]; more worked out versions are in Seibt 2004b, ch. 4; Seibt 2008, and, in particular, in Seibt 2009).

19. For example,  $P(\alpha)$  might represent the relationships in generic action contexts such as: removing the flat is part of changing a tire, loosening the lug nuts is part of removing the flat tire, turning the wrench is part of loosening the lug nuts. Alternatively, the partition can be taken to represent a scientific partition, e.g., an anatomical relationship: actin myofilament is part of a myofibril, which is part of a muscle fiber, which is part of a muscle. Note that the aim of LEM is to reconstruct the mereological inferential meaning of common sense and scientific kind terms; thus the structure of the partitions are given by relevant pieces of (common sense or scientific) discourse involving the kind terms in question.

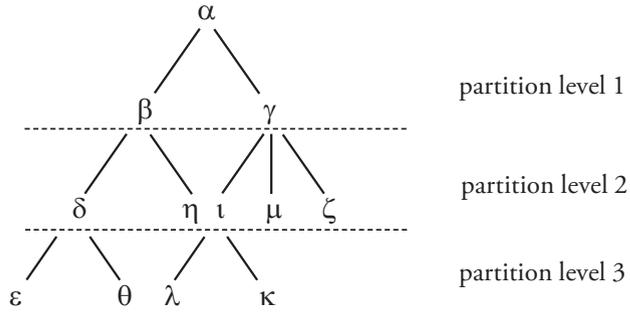


Figure 1: *Partition with three levels of nodes*

To introduce the notion of a partition and a partition level, we first define parthood resulting from the  $n$ -th iteration of ‘ $\hookrightarrow$ ’:

- (D1a)  $x \hookrightarrow_n y \leftrightarrow$  for all  $n$  in an ordered sequence of natural numbers of length  $N$ :  
 if  $n = 1$  then  $x \hookrightarrow_n y \leftrightarrow x \hookrightarrow y$   
 if  $n > 1$  then  $x \hookrightarrow_n y \leftrightarrow \exists z (x \hookrightarrow z \ \& \ z \hookrightarrow_{n-1} y)$

The relation ‘ $x \hookrightarrow_n y$ ’ (read: ‘ $x$  is  $n$ -part of  $y$ ’) indicates the level of mereological embedding of one item in another, but is always relative to a given partition.<sup>20</sup> For example, in figure 2 process  $\lambda$  is 3-part of  $\alpha$  and 2-part of  $\gamma$ . The  $n$ -parts of an entity  $\alpha$  are those entities  $y$  for which ‘ $y \hookrightarrow_n \alpha$ ’ holds, relative to a given partition:

- (D1b)  $x$  is  $n$ -part of  $y =_{df} x \hookrightarrow_n y$

Note that  $n$ -part is not a transitive relation. In order to define a part-relation that is transitive across partition levels down to the  $N$ -th level of the partition, we need to add a transitivity postulate (here omitted) for the relation ‘ $\hookrightarrow_{trans-N}$ ’.

- (D1c)  $x \hookrightarrow_{trans-N} y \leftrightarrow x \hookrightarrow y$  or  $x \hookrightarrow_n y$ , for any  $1 < n \leq N$

20. Not all partitions are finite; the identities of nodes of infinite partitions can be preserved by means of LEM’s relativized identity axioms, see below (Ax4a).

For a given entity  $\alpha$ , any entity  $y$  in the domain of the relation ' $y \text{ } \zeta_{i \leq n} \alpha$ ' belongs to the so-called *n-partition* of  $\alpha$ . In contrast, any entity  $y$  in the domain of ' $y \text{ } \zeta_n \alpha$ ' (the  $n$ -th iteration of ' $\zeta$ ') belongs to the so-called *partition-at-level- $n$*  of  $\alpha$ . That is, the *n-partition* of  $\alpha$  comprises entities at all partition levels up to and including level  $n$ , while the *partition-at-level- $n$*  of  $\alpha$  denotes entities at *one* level only. Items at level  $n$  of a partition  $P(x)$  are called the  $n$ -parts of  $x$  (see D1b), items below or above a certain 'threshold' can be easily referred to by quantifying over the '> $n$ -parts' or '< $k$ -parts' of  $x$  in  $P(x)$ , respectively. For example, in figure 1 above the '2-parts' of  $\alpha$  are  $\beta$  and  $\gamma$ , and the '>2-parts' of  $\alpha$  are  $\varepsilon$ ,  $\theta$ ,  $\lambda$ ,  $\kappa$ .

There are two sorts of mereological relationships that can be stated with the framework of (LEM): (i) universal relationships holding for any partition, either at levels, or across levels, or for certain thresholds (e.g., relative to certain domains of application); (ii) partition-relative relationships, holding only within a given partition or collection of partitions. Statements about partition-relative relationships may be made in two ways, however. They may relate to a certain partition level  $n$  of  $P(x)$  where  $x$  is the uppermost node; in some contexts it is important, however, to be able to address relationships relative to sub-partitions  $P(y)$  in  $P(x)$ , where  $y \text{ } \zeta_n x$  in  $P(x)$ . The expression ' $P(y \text{ } \zeta_n x)$ ' denotes a subpartition  $P$  of the partition of  $x$  that has  $y$  as its head node<sup>21</sup>; if  $z$  is a  $m$ -part of  $P(y \text{ } \zeta_n x)$ ,  $z$  is also a  $m+n$ -part of  $x$ . The simple index ' $n$ ' on the part-relation ' $\zeta_n$ ' indicates that a LEM-formula addresses an absolute partition level of a partition  $P(x)$ , while a composite index ' $\zeta_{m+n}$ ' indicates that a LEM-formula addresses a partition level of any sub-partition  $P(y \text{ } \zeta_n x)$ . Finally, the composite index ' $\zeta_{m(y)+n}$ ' is used to identify  $m$ -parts within the subpartition  $P(y \text{ } \zeta_n x)$  for some specific  $y$  that is  $n$ -part of  $P(x)$ . For example, in figure 1 above, the following relationships hold:  $\theta \text{ } \zeta_3 \alpha$ ;  $\theta, \lambda, \zeta_{2+1} \alpha$ ; and  $\theta \text{ } \zeta_{2(\beta)+1} \alpha$ .<sup>22</sup>

The distinction between relative and absolute partition levels in a partition also matters for a proper understanding of the anti-symmetry of the part-relation ' $\zeta$ '. Axioms (Ax2) and (Ax3) hold *at* any (absolute) partition level  $n$  in a partition  $P(z)$ , i.e., for *any given*  $n$  in a partition  $P(z)$  with  $n$ -parts  $x$  and  $y$ ,  $x \text{ } \zeta_n y \rightarrow \neg (y \text{ } \zeta_n x)$ , but this does not hold *across variable*  $n$ . This is ensured by (Ax4):

21. For the sake of simplicity I assume here that  $y$  occurs only once in  $P(x)$ .

22. One of the main tasks of (LEM) in application to processes is to investigate how absolute partition levels map into relative partition levels—i.e., how much of the mereological structure of a process is preserved once it becomes part of another.

$$(Ax4) \quad x \mathcal{C}_n z \ \& \ y \mathcal{C}_n z \ \rightarrow \ [\neg(x \mathcal{C}_{1(y)+n} z) \ \& \ \neg(y \mathcal{C}_{1(x)+n} z)]$$

If  $x$  and  $y$  are  $n$ -parts of the same partition  $P(z)$ ,  $x$  cannot be part of  $y$  in this partition, and vice versa. But note that (Ax4) leaves room for the possibility that  $x$  and  $y$  are parts of parts of ... parts of each other, i.e., mutually contain each other at deeper levels of respective sub-partitions of a main partition. For example, relative to a sub-partition  $P(\alpha \mathcal{C}_i \gamma)$  it may hold that  $\beta \mathcal{C}_3 \alpha$ , and relative to a sub-partition  $P(\beta \mathcal{C}_k \gamma)$  it may hold that  $\alpha \mathcal{C}_3 \beta$ . In a sense, then, for such a partition  $P(\gamma)$  one may claim that  $\alpha \mathcal{C}_3 \beta$  and  $\beta \mathcal{C}_3 \alpha$ , but only in the sense that ‘ $\mathcal{C}_3$ ’ in each case is *relative* to two different sub-partitions of  $P(\gamma)$ , not as a statement about the *absolute* partition levels of  $P(\gamma)$ .<sup>23</sup>

Perhaps the most important feature of (LEM) is that PPP, the identity principle, is defined across partitions but relative to a partition level, with the following two axiom schemas, as identity at a level  $n$  (see Ax5a) and identity relative to partition levels up to depth  $n$ , (see Ax5b, where ‘ $\mathcal{C}_{\leq n}$ ’ is used to abbreviate the restricted quantification over partition levels, read ‘for each  $m \leq n$ ’):

$$(Ax5a) \quad \forall z (z \mathcal{C}_n x \leftrightarrow z \mathcal{C}_n y) \leftrightarrow x =_n y$$

$$(Ax5b) \quad \forall z (z \mathcal{C}_{\leq n} x \leftrightarrow z \mathcal{C}_{\leq n} y) \leftrightarrow x =_{\leq n} y.$$

Two terms are coreferential (in common ontological parlance: ‘two entities are identical’) if and only if they have the same  $n$ -parts or  $\leq n$ -parts. These two identity principles account for the context-relativity of our judgments of sameness; in some contexts coarse-grained comparisons suffice, while others require an in-depth investigation. For example, in certain contexts it seems correct to identify the rotating of a metal disk with its heating up, or a performance of certain bodily movements with a greeting, or a singing with a production of noise, or a house with shelter, and such coarse-grained (subsumptive) identifications take their bearings from a cursory survey of what is involved or part of being a certain rotating of a disk, greeting, singing, or house. In other contexts, however, we

---

23. Again, that it might be desirable to have a formal framework where one can express, very generically, that two items are part of one another (e.g., trigger or modify one another) become most plausible if one applies (LEM) to processes (cf. my 2001, 2004a, b, 2009). Cortnoir (2010) provides additional motivations for accepting parthood-loops, but chooses to give up anti-symmetry altogether, while retaining transitivity.

will look more closely into deeper levels of the partitions of these items and find mereological incongruences in terms of which the items can be kept apart.<sup>24</sup>

Finally, to complete our sketch of the basic axioms of (LEM), given that ‘ $\hookrightarrow$ ’ is a non-transitive relation we can diversify the part-relation modeled by operating with ‘ $\hookrightarrow_n$ ’ instead of ‘ $\hookrightarrow$ ’ in various axioms and definitions.<sup>25</sup> For example, we can distinguish various degrees of overlap as well as three forms of ‘non-overlap’ among processes: *discreteness* reaches only into the first partition level, *disjointness* expresses non-overlap across levels, and *divergence* establishes non-overlap at a certain partition level. The supplementation principle in (LEM) is defined as a ‘shallow’ constraint in terms of discreteness (symbol: ‘ $\uparrow$ ’) rather than disjointness or divergence:

$$(A6a) \quad x \hookrightarrow y \rightarrow \exists z (z \uparrow x \ \& \ z \hookrightarrow y)$$

$$(A6b) \quad x \hookrightarrow_n y \rightarrow \exists z (z \uparrow_n x \ \& \ z \hookrightarrow_n y)^{26}$$

With such a weak version of the supplementation principle wholes may have ‘non-overlapping yet entangled’ parts, i.e., parts at partition level  $n$  that do not overlap at level  $n+1$  but share parts at some partition levels below the second level.

### 3. *Emergent parts of processes*

The preceding brief sketch of (LEM) is mainly intended as an illustration of *Leveled Mereology* as a general formal strategy; the idea of a partition-based mereology, the new notions of *parthood-at-level- $n$*  and *parthood-up-to-level- $n$* , and the relativization of the identity principle have a host of fruitful applications for the traditional problems of ontology (material constitution, persistence and change, individuation of events and actions). Moreover, there are many aspects of *Leveled Mereology* that provide useful connections between, on the one hand, the investigation of formal mereol-

24. This cognitive capacity of ‘zooming in and out,’ or moving down and up a partition to reason about more or less specific entities, we systematically exploit in many pragmatic contexts, from planning to excuses.

25. I am omitting here a discussion of certain constraints to be observed by such diversifications of axioms, for details see Seibt (201+b).

26. Discreteness at level  $n$ , i.e., ‘ $\uparrow_n$ ’ is defined as follows:  $x \uparrow_n z =_{\text{df}} x \hookrightarrow_n y$  and  $z \hookrightarrow_n y$  in  $P(y)$  and there is no  $w$  for which  $w \hookrightarrow_{1(x)+n} y$  and  $w \hookrightarrow_{1(z)+n} y$  (i.e.,  $w$  is 1-part in the sub-partitions  $P(x \hookrightarrow_n y)$  just in case it is not 1-part in the sub-partition of  $P(z \hookrightarrow_n y)$ , and vice versa).

ogy in ontology, and, on the other hand, the study of partitioning frames used in scientific explanations.<sup>27</sup>

I have explored the strategy of *Leveled Mereology* mainly in application to processes. In a recent publication Peter Simons writes: “Relatively little thought has however gone into the question whether we can simply adapt the mereology coming from mathematics and logic for processes” (2013: 161). I choose to read this statement as a claim about *numbers* of analytical ontologists engaged in such a project, and in this sense I can only agree.<sup>28</sup> The mereology of processes is a particularly difficult topic, in my view, for four reasons. First, we reason about processes (activities, events, developments, actions) in ways that defy ontological categorization in terms of the traditional category dualism of ‘concrete, determinate, particular, and individual’ entity versus ‘abstract, indeterminate, general’ entity—processes are best conceived as concrete, non-particular, indeterminate individuals. This new category I call *dynamics*.<sup>29</sup> Second, processes stand in non-transitive part-relations. Third, complex processes exhibit both sequential, circular, and multiply entangled structures (feedback, multiple entangled feedback); in order to accommodate non-sequential

---

27. Cf. Winther (2011, fn. 3), where this desideratum is mentioned. Winther’s account of scientific part-whole explanations, based on the claims that “there are multiple cross-cutting manners of abstracting a system into kinds of parts—i.e., there are multiple *partitioning frames*” (2011:397) and that “parts are *abstracted through partitioning frames closely linked to explanatory projects*” (400) fit directly with the constructive strategies of LEM, namely, (i) to begin with the most generic part-relation terms of which specific part-relations (functional, structural, or even more specific: morphological, physiological etc.) can be defined; (ii) to operate with a part-relation that—unlike transitive ‘part-of’—does not imply any implicit domain restrictions and allows, for instance, genes in the role of “structure-parts” and “activity-parts” (cf. *ibid.* 412); (iii) to associate terms with default partitions. I will need to leave the elaborations of these connections for another occasion, but note that *Leveled Mereology* also contains rules for operations on partitions (extension, reduction, rebranching, insertion, deletion) to model changes in the conceptual resources on which mereological reasoning is based, from simple predication to change of context to scientific progress.

28. Mereological relationships for processes also have been investigated in philosophy of science, especially in philosophy of chemistry (cf. e.g. Needham 2003) and biology (cf. e.g. Kaiser 2014; Winther 2012), but, as far as I can see, in analytical ontology the mereology of processes is largely uncharted; Mark Steen has explored logical relationships between stuffs and processes, Joseph Brenner developed a dialectical logic of processes (2008). I have addressed the issue from 1990 onwards, while working out the conceptual foundations for a new ontological category of non-particular concrete individuals called ‘dynamics’ or ‘general processes’ (cf. in particular 2004b ch. 4 and 5, but also 1990 ch. 5, 1995, 1997, 1999, 2000, 2001, 2004a, 2008, 2009, 2010).

29. In earlier terminological guises: ‘dynamic masses,’ ‘free processes,’ or ‘general processes’; for book-length arguments for this categorization (cf. in particular my 1990 and 2004b).

process architectures, a mereology of processes must allow for partition loops. Fourth, complex processes generate emergent processes that we may have reason to identify as ‘new’ parts of the complex process, or of one of its parts. It is this fourth ‘challenge’ for a mereology that I want to address in the following.<sup>30</sup>

In the longstanding debate about emergence, the keywords ‘self-organization,’ ‘self-maintenance,’ and ‘complexity’ have come to characterize systems exhibiting a behavior that can count as emergent in a fairly strong sense—the specifics of these behaviors are (physically) unpredictable even if they result from deterministic processes. Typical examples for such emergent behaviors are the development of convection cells (hurricanes, Benard cells), the joint movement of swarms, recurrent patterns in cellular automata, the differentiation of biological cell, a burning candle. Systems in which ‘self-organizing,’ ‘self-maintaining,’ or ‘complex’ behaviors occur have many different architectures; for present purposes I want to distinguish two sorts of emergence, simple and generative emergence. In cases of the first type, simple emergence, the emergent process does not itself causally influence the conditions for its continued existence—it is a product of an interaction of parts of the process system *S* that affects parts of *S* that are not identical with *S*. This type of emergence can be illustrated by C. Huygens’ virtual metronome. As Huygens noted, if several pendulum clocks are hung on one suspended beam, after a time the pendulums become synchronized and swing more evenly.<sup>31</sup> Let us describe the virtual metronome as a system of processes, using the folk-physical terms of common-sense reasoning. A relevant portion of the partition for such a system ( $\theta$ ) could look as follows, reducing the numbers of clocks to three:

---

30. As I hinted at above, partitions structured by the part-relation of LEM can contain loops and ‘mutual containments,’ provided these are not direct but mediated by a chain of part-of relations.

31. William Wimsatt cites Huygens’ “virtual metronome” as an illustration for what he calls “weak emergence”: “An emergent property is—roughly—a system’s property which is dependent on the mode of organization of the system, without being embedded in any part of the system per se” (Wimsatt 1994, 373). Wimsatt’s assessment that there is “nothing antireductionist” (ibid.) about the virtual metronome and similar cases of weak emergence depends on his assumption that the mode of organization is not part of the system.

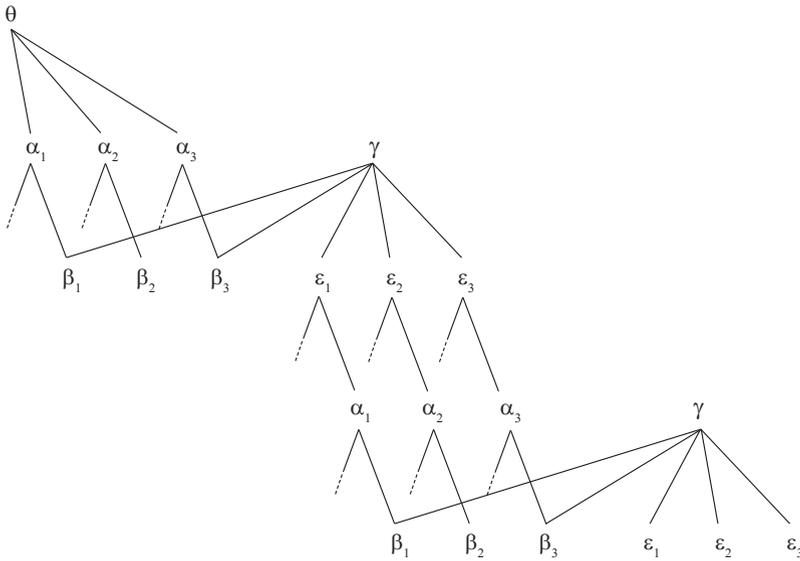


Figure 2: A (very simplified) analysis of virtual metronome  $\theta$ , in 'folk-physical' terms.  $\alpha_i$  = swinging of a pendulum in location  $i$ ;  $\beta_i$  = horizontal shifting of the beam in location  $i$  (i.e., transmitting horizontal momentum);  $\gamma$  = the interaction of the horizontal shiftings of the beam;  $\epsilon_i$  = dampening or accelerating of the pendulums (transmitting the momentum of the overall movement of the beam onto the oscillations) in location  $i$ .

Using ' $\zeta_n$ ' we can introduce in LEM the definitions of *sum-at-n* and *sum-up-to-n* relative to the  $n$ -th partition level of  $P(x)$ , which allows for the possibility of introducing terms for items that consist of the *sum-at-n* (or the *sum-up-to-n*) in  $P(x)$  and have additional parts that are not  $n$ -parts (or  $\leq n$ -parts) of  $x$ . Let us call such items 'interactions,' and the relevant additional parts they introduce 'interaction product terms.' For example, in Figure 2,  $\gamma$  is an interaction, consisting of a *sum-at-2* in  $P(\theta)$ , i.e., a sum of 2-parts of  $P(\theta)$ , and the interaction product term  $\delta$ ; there is a second occurrence of  $\gamma$  in  $P(\theta)$ , namely as *sum of 5-parts* of  $P(\theta)$ . The iteration of a part of process system is indicative of simple emergence with feedback.

The second type of emergence, generative emergence, is typical of self-maintaining systems—the emergent complex process, e.g., a burning candle, causally contributes to the continued occurrence of its component processes (far-from-equilibrium), e.g., the heating of the wax or the perco-

lating of the wax within the wick.<sup>32</sup> Represented with the tools of LEM, the partitions of systems with generative emergence display a distinctive iteration structure involving the headnode  $x$  of the partition  $P(x)$  within subpartitions of  $P(x)$ . To use again the case of self-maintenance, a process system  $S$  includes a self-maintaining process, just in case:

- (i) the partition  $P(S)$  contains at least one interaction  $y$ , with  $y$  being the sum of  $z_1 \dots z_k$  and  $S$ , where each of the  $z_1 \dots z_k$  is  $2 \leq n$ -part of  $P(S)$  for some given  $n$ ; and  $S$  is an interaction product term of  $y$ ;
- (ii)  $y \mathcal{C}_{m(S)+n} S$ , i.e., the interaction  $y$  is an  $m$ -part ( $2 \leq m$ ) of a subpartition with head node  $S$  that is embedded in  $P(S)$ .

Roughly speaking, the hallmark of partitions for process systems with productive emergence is that they have embedded partitions that are ‘self-similar’ in the sense that they contain structural repetitions involving the head-node of the embedded partition. It should be noted, though, that due to the fact that in LEM the identity of terms is defined relative to (i) the levels of (ii) a (relative/embedded or absolute) partition (see Axiom schema 5), we can choose whether system with generative emergence should be described as containing structural repetitions (choosing weak identity conditions for the terms of partition) or containing structural similarities (choosing strong identity conditions), so as to invite or avoid, respectively, idioms of ‘self’-maintenance.<sup>33</sup>

#### 4. Conclusion

I have argued in this paper for the use of a non-transitive generic part-relation that fits the most basic and indiscriminative usage of ‘part-of’ in the sense of ‘belonging with’ in common-sense reasoning. I suggested that the apparent naturalness of transitive reconstructions of ‘part-of’ is due to a combination of two factors. On the one hand, it reflects the historical provenance of formal mereology in the debate about the foundations of mathematics at the beginning of the 20<sup>th</sup> century; on the other hand, it

---

32. Cf. Bickhard (2003). The example of the burning candle is Bickhard’s illustration; it is discussed in some more detail in Seibt (2009).

33. Different degrees of self-similarity and types of emergence require various different types of adjustment on the ‘reach’ of ‘ $\mathcal{C}$ ’ in the axioms; for example, in partitions representing strong emergence (A2) and (A5) need to have the same reach. For details cf. Seibt (201+b).

corresponds with a longstanding theoretical habituation to think of individuals as particulars, i.e., entities that can be individuated in terms of their spatial location (at a time) alone. This mindset, which I call the ‘substance paradigm’ or the ‘myth of substance,’ facilitates the equivocation of a thing with the region that it occupies. Only if we allow for this equivocation, I argued, it will appear that spatial parthood, which is transitive on spatial regions, is also *per se* transitive.

Once we operate with a non-transitive part-relation, new theoretical strategies for mereology are opening up. I sketched one of these strategies, called *Leveled Mereology*, which I illustrated with a particular implementation LEM. LEM can be used for all domains but also, in particular, for generic (non-particular) individuals such as processes. I pointed at some of the potential merits of LEM for the modelling of parthood relationships on process architectures, in particular the possibility to work with partitions that have ‘loops,’ representing feedback or the entanglement of processes. In LEM one can quantify in a differentiated fashion over parts at a partition level, which makes it straightforward to represent interactions within in a partition; I explored this feature briefly for the representation of parts of processes that are emergent in two senses of this term.

The turn to *Leveled Mereology* signifies a rather momentuous reorientation, however, with respect to the aims of mereology. Classical mereology, in particular (CEM), corresponds to a view of mereology as a *logic* or *formal* theory of reasoning, where a few axioms and definitions suffice to construct inferences based on the main connective (part-of) alone, deriving a large number of theorems. This conception is somewhat compromised already by intensional mereologies, where ‘material’ aspects enter by way of identity principles for certain domains and additional axiom sets for specifications of the main connective (functional part, construction part etc.). The strategy of *Leveled Mereology* in effect takes mereological reasoning to be about pairs  $\langle x, P(x) \rangle$  of terms and partitions, and thus leads even farther away from logic and into ‘knowledge representation.’ As long as the partitions of LEM operate merely with the generic relation of ‘part-of,’ few theorems can be derived; once more specific types of parthood (material, functional, design, construction, repair parts etc.) are defined in terms of this generic relation and for a certain region of the ontological domain, a host of specific inference principles can be stated. In short, the interest of LEM lies in the clarification and discussion of claims about mereological structures (especially for processes). LEM contains a few principles that

characterize the generic notion of ‘part’ but is designed as a general formal framework for the representation of regional (context-sensitive) mereological knowledge and regional principles of mereological inference.

#### REFERENCES

- Bickhard, Mark H. 2003: “Process and Emergence: Normative Function and Representation”. In: Johanna Seibt (ed.), *Process Theories*. Netherlands: Springer, 121–155.
- Brenner, Joseph E. 2008: *Logic in Reality*. Netherlands: Springer.
- Cotnoir, Aaron J 2010: “Anti-Symmetry and Non-Extensional Mereology”. *The Philosophical Quarterly* 60 (239), 396–405.
- Cruse, D. Alan 1979: “On the Transitivity of the Part-Whole Relation”. *Journal for Linguistics* 15(1), 29–38.
- Dowty, David 1977: “Toward a Semantic Analysis of Verb Aspect and the English ‘Imperfective’ Progressive”. *Linguistics and Philosophy* 1(1), 45–78.
- Doepke, Frederick 1982: “Spatially Coinciding Objects”. *Ratio* 10(24), 45–60.
- Guizzardi, Giancarlo 2009: “The Problem of Transitivity of Part-Whole Relations in Conceptual Modeling Revisited”. In: *Advanced Information Systems Engineering*. Berlin Heidelberg: Springer, 94–109.
- Mourelatos, Alexander 1978: “Events, Processes, and States”. *Linguistics and Philosophy* 2(3), 415–434.
- Needham, Paul 2004: “Continuants and Processes in Macroscopic Chemistry”. *Axiomathes* 14(1), 237–265.
- Puntel, Lorenz B. 2010: *Structure and Being: A Theoretical Framework for a Systematic Philosophy*. University Park, PA: Penn State Press.
- Sanford, David H. 1993: “The Problem of the Many, Many Composition Questions, and Naive Mereology”. *Noûs* 27(2), 219–228.
- Seibt, Johanna 1995: „Individuen als Prozesse: Zur prozess-ontologischen Revision des Substanzparadigmas“. *Logos* 2, 352–384.
- 1990: *Towards Process-Ontology: A Critical Study of Substance-Ontological Premises*. Ph.D. Dissertation at the University of Pittsburgh. Michigan: UMI-Press. To appear in a revised version as: *The Myth of Substance*. Berlin: DeGruyter.
- 1996: “Existence in Time: From Substance to Process”. In: Jan Faye, Uwe Scheffler & Max Urs (eds.), *Perspectives on Time. Boston Studies in Philosophy of Science*. Dordrecht: Kluwer, 143–182.
- 1999: “Dinge als Prozesse”. In: Rafael Hüntelmann & Erwin Tegtmeier (eds.), *Neue Ontologie und Neue Metaphysik*. Köln: Academia Verlag, 11–41.

- Seibt, Johanna 2000: “The Dynamic Constitution of Things”. In: Jan Faye, Uwe Scheffler & Max Urchs (eds.), *Things, Facts and Events. Poznan Studies in Philosophy of Science*. Amsterdam: Rodopi, 241–279.
- 2001: “Formal Process Ontology”. In: Chris Welty & Barry Smith (eds.), *Formal Ontology in Information Systems: Collected Papers from the Second International Conference*. Ogunquit: ACM Press, 333–345.
- (Ed.) 2003: *Process Theories—Crossdisciplinary Studies in Dynamic Categories*. Dordrecht: Kluwer.
- 2004: “Free Process Theory: Towards a Typology of Processes”. *Axiomathes* 14(1), 23–57.
- 2005a: *General Processes—A Study in Ontological Category Construction*. Habilitationsschrift at the University of Konstanz, Germany.
- 2005b: “Der Mythos der Substanz”. In: Käthe Trettin (ed.), *Substanz—Neue Überlegungen zu einer klassischen Kategorie des Seienden*. Frankfurt: Klostermann, 197–229.
- 2008: “Beyond Endurance and Perdurant: Recurrent Dynamics”. In Christian Kanzian (ed.), *Persistence*. Frankfurt: Ontos Verlag, 121–153.
- 2009: “Forms of Emergent Interaction in General Process Theory”. *Synthese* 166 (3), 479–512.
- 2012: “Process Philosophy”. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2013/entries/process-philosophy/>>.
- 2014: “Transitivity”. In: Hans Burkhardt, Johanna Seibt & Guido Imaguire (eds.), *Handbook of Mereology* (forthcoming). München: Philosophia Verlag.
- 201+a: “Are There Transitive Part-Relations?” (under submission).
- 201+b: “Leveled Mereology” (under submission).
- Simons, Peter & Dement, Charles 1996: “Aspects of the Mereology of Artifacts”. In: Roberto Poli & Peter Simons (eds.), *Formal Ontology*. Dordrecht: Kluwer, 255–276.
- Simons, Peter 1987: *Parts: A Study in Ontology*. Oxford: Oxford University Press.
- 2000: “How to Exist at a Time When You Have No Temporal Parts”. *The Monist* (83)3, 419–436.
- 2013: “Varieties of Parthood: Ontology Learns from Engineering”. In: *Philosophy and Engineering: Reflections on Practice, Principles and Process*. Netherlands: Springer, 151–163.
- Varzi, Achille 2006: “A Note on the Transitivity of Parthood”. *Applied Ontology* 1 (2), 141–146.
- Vieu, Laure 2006: “On the Transitivity of Functional Parthood”. *Applied Ontology* 1 (2), 147–155.

- Wimsatt, William C. 1994: "The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets". *Canadian Journal of Philosophy* 24 (1), 207–274.
- Winston, Morton, Robert Chaffin & Douglas Herrmann 1987: "A Taxonomy of Part-Whole Relations". *Cognitive Science* 11 (4), 417–444.
- Winther, Rasmus G. 2011: "Part-whole Science". *Synthese* 178 (3), 397–427.

## NOTHING OVER AND ABOVE\*

Theodore SIDER  
Cornell University

### *Summary*

The slogan “the whole is nothing over and above the parts” animates many theories of parthood and is central to our ordinary conception. I examine the slogan’s relationship to classical mereology, Kit Fine’s “operationalist” account of “mere sums”, composition as identity, and unrestricted composition.

“The whole is nothing over and above the parts”—this slogan animates many theories of parthood, is intuitively compelling, and is arguably central to our ordinary conception. Yet it seems to make no sense. As I understand it, the slogan says that an exceedingly intimate relationship holds between a whole and its parts: in some sense the whole is *no different from* its parts. But how can that be? I am just one thing; my head, arms, legs, and torso are more than one in number; so how can I be “nothing over and above” or “no different from” them?

The slogan is admittedly vague. But there are various precise theses purporting to capture the slogan’s spirit whose truth we can meaningfully debate. The murky question of whether the theses really capture the spirit will remain—and I do not mean to downplay the importance of this question—but at least the murk will be contained. First we will consider a boring (though perhaps ultimately the best) precise rendering of the slogan’s spirit, which is simply that classical mereology is true. We’ll then discuss some more exciting ways to precisify the slogan, and conclude by asking whether the exciting ways promise anything better.

---

\* Like everyone interested in questions about parthood, especially those with a formal side, I am deeply indebted and grateful to Peter Simons for his pioneering contributions. Thanks also to the participants in the summer 2012 Metaphysical Mayhem at Rutgers, where some of this material was originally presented, and to Karen Bennett, Phillip Bricker, Ross Cameron, Kit Fine, Kris McDaniel, Jonathan Schaffer, and Jason Turner.

‘Nothing over and above’ is a flexible piece of philosophical rhetoric, applicable across a variety of situations and to entities of various categories. A fact might be said to be nothing over and above another when it is necessitated or grounded by the latter fact; a property might be said to be nothing over and above another when it is realized by the latter property; one might say that wholes are qualitatively nothing over and above their parts meaning that composite objects possess no “emergent” properties; and so on.<sup>1</sup> But the slogan to be explored here concerns, it would seem, a narrower sense of ‘nothing over and above’: that of a certain especially intimate ontic relation between a thing and its parts. Though I will be open to attempts to explicate this narrower sense in terms of other sorts of “nothing over and above”, the narrower sense itself is our target.

### 1. *Classical mereology*

Classical mereology is a formal theory of parts and wholes.<sup>2</sup> A typical presentation: take ‘ $x < y$ ’ as a primitive predicate for parthood, define overlap (“ $Oxy$ ”) as sharing a part in common, discreteness (“ $Dxy$ ”) as nonoverlap, proper parthood (“ $x \ll y$ ”) as parthood without identity, and fusion as follows:<sup>3</sup>

$x \text{ Fu } S$  (“ $x$  is a fusion of set  $S$ ”) =<sub>df</sub> each member of  $S$  is part of  $x$ , and each part of  $x$  overlaps some member of  $S$

As axioms, assume that  $<$  is reflexive, transitive, antisymmetric, and also obeys:<sup>4</sup>

**Weak supplementation** If  $x \ll y$ , then some part of  $y$  is discrete from  $x$

**Unrestricted composition:** For any nonempty  $S$  there exists a fusion of  $S$

1. The broad sense may be akin to Karen Bennett’s notion of building. Bennett also provides an interesting way to take the slogan, albeit one that is very distant from those considered here: positing nonfundamental entities does not count against the simplicity of a theory (2015, chapter 7).

2. See Hovda (2009); Simons (1987).

3. When  $x \text{ Fu } \{y_1, \dots\}$  we may say informally that  $x$  is a fusion of  $y_1, \dots$ .

4. This axiom set is not minimal: reflexivity and antisymmetry can, for instance, be eliminated (Hovda, 2009). Since we will be interested in certain subsystems of classical mereology, it’s best to retain the redundancy.

One way to precisify our slogan would be to say that each theorem of classical mereology—or perhaps, each theorem in a certain chosen subsystem of classical mereology—is true. Why would this count as a precisification of the slogan, a way of capturing the vague idea that the whole is nothing over and above the parts? Because, as we will see, one can think of the axioms of classical mereology—some of them, anyway, and perhaps all—as being in some sense underwritten by the slogan, much as (some of) the axioms of Zermelo-Frankel set theory are often regarded as being underwritten by the intuitive iterative conception of set.<sup>5</sup>

Showing axioms in a formal system to be “underwritten” by a vague slogan is an inherently nonrigorous process, but part of the process can be made rigorous. The *nonrigorous* part is a phase of “regimentation”, in which we regiment the concept of “nothing over and above”, lay down precise principles on how the regimentation behaves that seem to be inspired by its intended—though unclear and perhaps incoherent!—content, and precisely formulate the slogan based on the regimentation. This phase is nonrigorous because there can be no formal proof that the laid-down principles and precise formulation are faithful to the intended intuitive content of “nothing over and above” and the slogan. (Compare the relationship between the informal notion of computability and its various formalizations.) But once this is done, the remaining “derivation” phase is perfectly rigorous: we can derive axioms of mereology from the regimented slogan plus the laid-down principles. (Compare what Boolos (1971) did for the iterative conception of set.) I will show how to do this for the axioms of reflexivity, transitivity, antisymmetry, and weak supplementation.<sup>6</sup> (We will discuss unrestricted composition in section 4.)

Let’s begin with the regimentation phase. First we regiment ‘nothing over and above’ as a two place predicate  $\approx$ :<sup>7</sup>

$x \approx S$      “object  $x$  is nothing over and above the members of set  $S$ ”

Next we lay down these principles governing  $\approx$ :

$\approx$ -**Reflexivity**  $x \approx \{x\}$

---

5. Thanks to Jonathan Schaffer for this analogy.

6. The arguments are analogous to those in Sider (2007, 3.2); see also Sider (2007, 4.2) and Lewis (1991, 3.6).

7. Though  $\approx$  is a relation to sets, nothing-over-and-above is not. The intended gloss of “ $x \approx \{y_1, \dots\}$ ” is “ $x$  is nothing over and above  $y_1, \dots$ ”, not “ $x$  is nothing over and above  $\{y_1, \dots\}$ ”.

**Cut** If  $x \approx S \cup \{y\}$  and  $y \approx T$  then  $x \approx S \cup T$

**$\approx$ -Uniqueness** if  $x \approx S$  and  $y \approx S$  then  $x = y$

These are natural assumptions. For instance, Cut tells us that if the army is nothing over and above the soldiers, and if each soldier is nothing over and above her molecules, then the army is nothing over and above all the molecules of all its soldiers; and  $\approx$ -Uniqueness tells us that a statue and a lump of clay cannot both be nothing over and above—no different from—the very same molecules. Finally, we formulate the slogan based on the regimentation:

**Slogan** For any object  $x$  and set  $S$ ,  $x \text{ Fu } S$  if and only if  $x \approx S$

Note three features of this formulation. First, we have not restricted the slogan to certain privileged decompositions of wholes (such as to proper parts or to atoms); we have interpreted it as saying that for *each* decomposition of the whole, the whole is nothing over and above the members of that decomposition.<sup>8</sup> Second, we have interpreted “whole” in the slogan as fusion in the sense of Fu. And third, we have interpreted the slogan “biconditionally”: as not merely implying that if  $x \text{ Fu } S$  then  $x \approx S$ , but as also implying the converse.<sup>9</sup> “Nothing over and above” must be understood as in the beginning of the paper if the converse is to be plausible.<sup>10</sup> There I gave the alternate gloss: “no different from”, which implies “no less than” as well as “no more than”. Thus, where  $x_1, x_2, \dots$  are all my subatomic particles, a mere proper part of me such as my hand does not count as being nothing over and above  $x_1, x_2, \dots$  in the present sense.

Now for the derivation phase: reflexivity, transitivity, antisymmetry, and weak supplementation follow from the regimented slogan and the principles governing  $\approx$ :

*Reflexivity*: for any object  $x$ , by  $\approx$ -Reflexivity,  $x \approx \{x\}$ ; by the slogan,  $x \text{ Fu } \{x\}$ ; by the definition of Fu,  $x < x$ .

*Transitivity*: reflexivity (which was just derived from the slogan) and the definition of Fu imply:

---

8. Thus in particular we have not interpreted the slogan as saying (merely) that the whole is nothing over and above the collection consisting of *all* its parts.

9. Compare the distinction between “strong” and “superstrong” composition as identity in Sider (2007).

10. Likewise for  $\approx$ -Uniqueness.

(+) if  $a < b$  then  $b \text{ Fu}\{a, b\}$

So now, assume  $x < y$  and  $y < z$ . By (+),  $y \text{ Fu}\{x, y\}$  and  $z \text{ Fu}\{y, z\}$ ; by the slogan,  $y \approx \{x, y\}$  and  $z \approx \{y, z\}$ ; by Cut,  $z \approx \{x, y, z\}$ ; by the slogan,  $z \text{ Fu}\{x, y, z\}$ ; by the definition of Fu,  $x < z$ .

$\approx$ -Uniqueness plus the slogan imply the principle of uniqueness of fusions:

**Uniqueness** if  $x \text{ Fu } S$  and  $y \text{ Fu } S$  then  $x = y$

(Uniqueness is perhaps most clearly underwritten by the slogan of all the principles of mereology.) Uniqueness, (+), and reflexivity then imply the remaining two axioms:

*Antisymmetry*: if  $x < y$  and  $y < x$  then by (+),  $x \text{ Fu}\{x, y\}$  and  $y \text{ Fu}\{x, y\}$ , and so by uniqueness,  $x = y$ .

*Weak supplementation*: suppose for reductio that i)  $x < y$ , ii)  $x \neq y$ , and iii) each part of  $y$  overlaps  $x$ . Given i) and iii), by the definition of Fu,  $y \text{ Fu}\{x\}$ . Given reflexivity and the definition of Fu,  $x \text{ Fu}\{x\}$ ; so by uniqueness,  $x = y$ , which contradicts ii).

There is a sense, then, in which the slogan that the whole is nothing over and above the parts underwrites reflexivity, transitivity, antisymmetry, and weak supplementation. But there is a wrinkle: my derivations of the axioms from the regimented slogan are sensitive to certain details of my chosen formulation of mereology.

For example, my chosen notion of fusion was Fu; but other formulations of classical mereology sometimes use other notions of fusion, such as:<sup>11</sup>

$x \text{ Fu}^* S =_{\text{df}}$  for all  $y$ ,  $y$  overlaps  $x$  iff  $y$  overlaps at least one member of  $S$

$x \text{ Fu}^{\text{lub}} S =_{\text{df}}$  each member of  $S$  is part of  $x$  and  $x$  is part of everything that contains each member of  $S$  as part (i.e.,  $x$  is a least upper bound of  $S$ )

This raises a concern. As I mentioned earlier, I formulated the slogan as concerning fusion *in the Fu sense of 'fusion'*. Accordingly, when drawing out consequences of the slogan, I made use of the definition of Fu. But suppose the slogan were instead understood in the sense of  $\text{Fu}^*$  or  $\text{Fu}^{\text{lub}}$

---

11. See Simons (1987, 35).

(that is, as saying that  $x \text{ Fu}^* S$  iff  $x \approx S$ , or that  $x \text{ Fu}^{\text{lub}} S$  iff  $x \approx S$ ). In that case the slogan's implications would not concern fusion in the Fu sense, and my arguments would not immediately apply. At best, analogous arguments, appealing to the definition of the replacement notion of fusion, could be constructed.

In many contexts it does not matter which of the three notions of fusion one uses, since given the entirety of classical mereology the three notions are equivalent.<sup>12</sup> But in the present context we are trying to *justify* classical mereology (by deriving certain axioms from the slogan), so the equivalence of the three notions cannot simply be assumed. There is, therefore, a question of whether one can still argue from the slogan to the axioms when one interprets the slogan in the sense of  $\text{Fu}^*$  or  $\text{Fu}^{\text{lub}}$ .

The crucial move in my argument for reflexivity was from  $x \text{ Fu}\{x\}$  to  $x < x$ ; this was justified because the definition of Fu logically implies that each member of  $S$  is part of any Fu-fusion of  $S$ . Since the definition of  $\text{Fu}^{\text{lub}}$  shares this feature—it logically implies that each member of  $S$  is part of any  $\text{Fu}^{\text{lub}}$ -fusion of  $S$ —the argument for reflexivity still goes through if the slogan is interpreted using  $\text{Fu}^{\text{lub}}$ . But the definition of  $\text{Fu}^*$  does not on its own logically imply that members of  $S$  are parts of any Fu-fusion of  $S$  (and there is no other available justification of the crucial move in the argument for reflexivity). The claim that  $x \text{ Fu}^* S$  says merely that for all  $z$ ,  $z$  overlaps  $x$  iff  $z$  overlaps some member of  $S$ . To be sure, this does imply, given further principles from classical mereology, that each member of  $S$  is part of  $x$ ; but in the present context those further principles would need first to be derived from the slogan, and I see no way to do that.

The situation with transitivity and antisymmetry is parallel. The arguments for those axioms relied on principle (+), which is a logical consequence of the definition of Fu and reflexivity. The corresponding principle for  $\text{Fu}^{\text{lub}}$  is also a logical consequence of the definition of  $\text{Fu}^{\text{lub}}$  plus reflexivity; but the corresponding principle for  $\text{Fu}^*$  does not follow from the definition of  $\text{Fu}^*$  without the use of further principles of classical mereology.

Thus the case for reflexivity, transitivity, and antisymmetry on the basis of the slogan requires that the slogan be true either in the sense of Fu or  $\text{Fu}^{\text{lub}}$  (or both). (Let us set aside the possibility of taking the slogan in the sense of some fourth notion of fusion.) The case would fail if the

---

12. See Hovda (2009); the axiomatization differs depending on the chosen notion of fusion.

slogan were true *only* in the sense of  $Fu^*$ . Now, it may be argued that  $Fu^*$  is not an intuitive notion of fusion, precisely because its definition does not guarantee that a fusion of some things must contain those things as parts. If this is right, then assuming that the slogan expresses an intuitive insight into the nature of composition and parthood, one would expect it to be true under some natural, intuitive notion of fusion, and hence under either  $Fu$  or  $Fu^{lub}$  (or both), in which case the derivations of reflexivity, transitivity, and antisymmetry go through. The case for those three axioms, then, arguably remains robust.

The status of weak supplementation, though, is less clear.  $Fu^{lub}$  is surely a reasonable notion of fusion; and the argument from the slogan to weak supplementation apparently fails if the slogan is understood in terms of  $Fu^{lub}$ . The argument began by deriving the uniqueness of fusions from the slogan, and then derived weak supplementation from uniqueness of fusions. But if the slogan concerns  $Fu^{lub}$  then the first stage delivers merely uniqueness of fusions in the sense of  $Fu^{lub}$ , which does not imply weak supplementation.<sup>13</sup>

The slogan, then, implies weak supplementation under one natural notion of fusion but not under another such notion. So does the slogan underwrite the axiom? Should a friend of the slogan accept it under both readings, or just one; and if the latter, which one? Perhaps the slogan is too unclear to deliver a verdict.

A related problem concerns different choices of primitive notions. I said above that my arguments from the slogan to reflexivity, transitivity, and antisymmetry do not go through when the slogan is understood in terms of  $Fu^*$ . However, that was under the assumption that  $<$  is an undefined notion. Classical mereology is sometimes formulated with overlap as the primitive notion, and with “ $x < y$ ” defined as “ $\forall z(zOx \rightarrow zOy)$ ”. Reflexivity and transitivity of parthood are logical consequences of this definition, and so one might regard them as not needing justification by the slogan.<sup>14</sup> Moreover, the slogan, understood in terms of  $Fu^*$ , now implies

---

13. In a model containing just two things, an object  $b$  and a single proper part,  $a$ , of  $b$  (each of which is part of itself), weak supplementation fails while uniqueness in the  $Fu^{lub}$  sense holds (as do reflexivity, transitivity, and antisymmetry). The crucial thing is that although both  $b$  and  $a$  count as  $Fu$ -fusions of  $\{a\}$ ,  $b$  is not a least upper bound of  $\{a\}$  since it's not part of  $a$ , which is an upper bound of  $\{a\}$ . (Compare Hovda's (2009, 67) remark that the point of weak supplementation is to insure that  $Fu$ -fusions are least upper bounds.)

14. In this section I have been assuming that definitions, such as the definition of  $Fu$  in terms of  $<$ , do not need to be justified by the slogan; this, though, might be questioned.

antisymmetry.<sup>15</sup> Of course, once we shift from  $<$  to  $O$  as the primitive notion of mereology, new axioms are needed, and their justification from the slogan will need to be considered in turn. The point is just that, as before, the amount of classical mereology implied by the slogan depends on how exactly we understand the notion of whole—that is, fusion—in the slogan. As we saw before, it matters how fusion is to be defined in terms of parthood, and we now see that it also matters, even given a fixed definition of fusion in terms of parthood, whether parthood is to be primitive or defined.

This section has been a little disappointing. My derivation of axioms of mereology was distressingly sensitive to subtle matters of how the slogan is to be formalized (particularly in the case of weak supplementation). Moreover, even setting this concern aside, my formalization of the slogan might be regarded as stretching the slogan’s spirit. It would be nicer to find a precise thesis that is clearly in the spirit of the slogan, and from which classical mereology, or a good portion of it anyway, could be robustly derived.

## 2. *Fine on mere sums*

Kit Fine’s fascinating article “Towards a theory of part” introduces a precise sense in which the composite objects of classical mereology are nothing over and above their parts, or are “mere sums”, as he puts it. Let us examine this idea, and its relation to the unresolved issues of the previous section.

The main point of Fine’s article is to defend two claims about parthood: *pluralism* and *operationalism*. According to pluralism, there are many different kinds of parthood relations. Each of the following relations is a kind or species of parthood, according to Fine: the relation of parthood in classical mereology, the relation of membership in set theory, the relation between sequences and their members, and the relation between propositions and their constituents.<sup>16</sup>

According to operationalism, relations of parthood are to be defined in terms of operations of composition, rather than the other way around. An operation is like a function: it yields an output if you give it inputs. Gram-

15. If  $x < y$  and  $y < x$  then by the definition of  $<$ ,  $\forall z(zOx \leftrightarrow zOy)$ . This implies  $\forall z(zOx \leftrightarrow (zOx \vee zOy))$ , so we have  $x Fu\{x, y\}$ . Similarly, we have  $y Fu\{x, y\}$ . So by uniqueness of  $Fu^*$ -fusions (which is justified by the slogan, now interpreted using  $Fu^*$ ),  $x = y$ .

16. In each of the latter three cases it is really the ancestral of the relation in question that is a species of parthood. On mereological pluralism see also McDaniel (2004, 2009).

matically, expressions for operations combine with terms to form terms. So, to take an arithmetic example, the terms ‘2’ and ‘3’ combine with the operation expression ‘ $\times$ ’ to yield the term ‘ $2 \times 3$ ’; similarly, ‘1’ and ‘ $2 \times 3$ ’ combine with the operation expression ‘+’ to yield the term ‘ $1 + (2 \times 3)$ ’. If  $\Sigma$  is an operation of composition taking operands  $y_1, y_2, \dots$  to the object  $x$  that they compose, we may write:  $x = \Sigma(y_1, y_2, \dots)$ . A composition operation is an operation whereby wholes are generated out of parts. Given an operation  $\Sigma$  of composition, an associated relation of parthood may be defined. First we define the operation’s associated relation of constituency, or “direct parthood”:

**Constituent**  $x$  is a constituent of  $y =_{\text{df}} y$  is the result of applying  $\Sigma$  to  $x$ , perhaps together with some other objects (i.e.,  $y = \Sigma(\dots x \dots)$ )

Parthood may then be defined as the ancestral of constituency:

**Parthood**  $x < y =_{\text{df}} x$  is a constituent of  $y$ , or a constituent of a constituent of  $y$ , or...

Thus Fine reverses the usual order of definition in classical mereology, in which a predicate for parthood ( $<$ ) is taken as primitive and then is used to define an operation of composition (fusion).

Given Fine’s operationalism, his pluralism is in the first instance a pluralism about composition operations. (This then induces pluralism about parthood relations.) Composition operations for Fine include an operation of mereological summation (“ $\Sigma$ ” henceforth), the “set-builder” operation, the “sequence-builder” operation, and so forth.<sup>17</sup>

Now, Fine does not accept the nothing-over-and-above slogan for all forms of composition, but he does accept it for mereological summation:

There are two aspects of the notion of whole that have often been implicit in the recent development of mereology. The first, more formal aspect is that a whole is a ‘mere sum’. It is nothing over and above its parts—or perhaps we should say, more cautiously, that it is nothing over and above its parts except insofar as it is one object rather than many. (2010, 572)

He then goes on to articulate the idea that mere sums are “nothing over and above” their parts in terms of a constraint on  $\Sigma$ :

---

17. Compare Karen Bennett’s (2011; 2015) pluralism about “building relations”.

**Summative identity:** Any regular identity condition is true

where an identity condition is a formula of the form ' $s = t$ ', where  $s$  and  $t$  are terms in  $\Sigma$ , a term in  $\Sigma$  is any term constructed solely from variables and  $\Sigma$ , a formula ' $s = t$ ' is regular iff the same variables occur in  $s$  and  $t$ , and a formula is true iff it is true for all values of its free variables. So, for instance,  $\Sigma(x, \Sigma(y, z)) = \Sigma(y, \Sigma(x, z))$  is regular, and so is true for all values of its variables.

Fine goes on to derive the condition of summative identity from some more basic constraints on  $\Sigma$ :

**Absorption:**  $\Sigma(\dots, x, x, \dots, y, y, \dots) = \Sigma(\dots, x, \dots, y, \dots)$

**Collapse:**  $\Sigma(x) = x$

**Leveling:**  $\Sigma(\dots, \Sigma(x, y, z, \dots), \dots, \Sigma(u, v, w, \dots), \dots) = \Sigma(\dots, x, y, z, \dots, u, v, w, \dots)$

**Permutation:**  $\Sigma(x, y, z, \dots) = \Sigma(y, z, x, \dots)$   
(and similarly for all other permutations)

Since each of these four constraints is a regular identity condition, each follows from summative identity. Thus summative identity is equivalent to the four taken together.

Now, Fine says that absorption, collapse, leveling, and permutation “constitute an analysis of the notion of mere sum” (574); and he says that:

[Summative identity] gives formal expression to the idea that wholes built up from the same parts should be the same, and this is something that appears to be constitutive of our intuitive conception of a mere sum as nothing over and above its parts ... Thus philosophical reflection on the notion of mere sum is able to provide us with a simple and natural characterization of classical mereology. (572)

So it is natural to wonder whether summative identity gives us what we are after in this paper: a precise thesis capturing the slogan. And it is also natural to wonder how much of classical mereology is implied by summative identity.

Summative identity does indeed seem to be underwritten by the slogan. Not only is this the case on an intuitive level, summative identity is implied

(modulo the existence of  $\Sigma$ -composites) by an appropriate regimentation of the slogan in terms of  $\Sigma$ , plus the assumptions about  $\approx$  laid down in the previous section.<sup>18</sup> (This is unsurprising given the close correspondence of Cut to Leveling,  $\approx$ -reflexivity to Collapse, and the set-theoretic regimentation of  $\approx$  to Absorption and Permutation.) Moreover, there is no puzzle like that at the end of the previous section of which notion of fusion is to be used in interpreting the slogan: it's natural for an operationalist to insist on taking the slogan as concerning  $\Sigma$ .

But it is less clear whether summative identity captures the entirety of the slogan. First note that it does not imply all of classical mereology (not that Fine says it does). It does imply reflexivity, transitivity, and antisymmetry.<sup>19</sup> But it does not imply weak supplementation.<sup>20</sup> Consider a model

---

18. Begin by using  $\Sigma$  as a predicate rather than a term-forming operator: let  $x\Sigma(y_1\dots)$  mean that  $x$  is a  $\Sigma$ -sum of  $y_1, \dots$ , with no assumption of uniqueness. Assume an infinitary language allowing infinite lists of variables (and other terms) of arbitrarily high cardinality. Take the slogan as the schema " $x\Sigma(y_1\dots)$  if and only if  $x \approx \{y_1\dots\}$ ". Uniqueness of  $\Sigma$ -sums then follows from the slogan and  $\approx$ -uniqueness. So we may henceforth use  $\Sigma$  as a term-forming operator (modulo the assumption of unrestricted  $\Sigma$ -composition, which we here simply assume), and reformulate the slogan thus:  $\Sigma(y_1\dots) \approx \{y_1\dots\}$ . Now for Fine's four principles. Absorption: by the slogan,  $\Sigma(A, x, x, B, y, y, C) \approx \{A, x, x, B, y, y, C\}$  and  $\Sigma(A, x, B, y, C) \approx \{A, x, B, y, C\}$ ; but  $\{A, x, x, B, y, y, C\} = \{A, x, B, y, C\}$ , and so by  $\approx$ -uniqueness,  $\Sigma(A, x, x, B, y, y, C) = \Sigma(A, x, B, y, C)$ . (Uppercase variables  $A, B, \dots$  are schematic for arbitrary lists of terms.) Permutation is similar. Collapse: by the slogan,  $\Sigma(x) \approx \{x\}$ ; by  $\approx$ -reflexivity,  $x \approx \{x\}$ ; by  $\approx$ -uniqueness,  $\Sigma(x) = x$ . Leveling: we'll establish  $\Sigma(A, \Sigma(B), C) = \Sigma(A, B, C)$  (the proof of the general case, i.e leveling multiple  $\Sigma$  terms simultaneously, is parallel). Given absorption and permutation, it suffices to establish  $\Sigma(A, \Sigma(B)) = \Sigma(A, B)$ . By the slogan,  $\Sigma(A, \Sigma(B)) \approx \{A, \Sigma(B)\}$  and  $\Sigma(B) \approx \{B\}$ ; by cut,  $\Sigma(A, \Sigma(B)) \approx \{A, B\}$ ; by the slogan,  $\Sigma(A, B) \approx \{A, B\}$ ; by  $\approx$ -Uniqueness,  $\Sigma(A, \Sigma(B)) = \Sigma(A, B)$ .

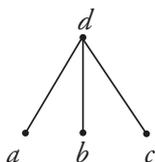
19. Reflexivity follows from Collapse. Transitivity follows directly from Fine's general definition of parthood in terms of composition, without any special assumptions involving  $\Sigma$ . However, Fine points out that a more substantive question is whether the relation of being a component is transitive; and given Leveling, it is. (Suppose  $x$  is a component of  $y$  and  $y$  is a component of  $z$ . Then  $y = \Sigma(A, x)$  and  $z = \Sigma(B, y)$ , for some  $A$  and  $B$ ; whence  $z = \Sigma(B, \Sigma(A, x))$ , which by leveling is  $\Sigma(B, A, x)$ , so  $x$  is a component of  $z$ .) To establish antisymmetry, let  $x < y$  and  $y < x$ . Then (a)  $y = \Sigma(A, x)$  and (b)  $x = \Sigma(B, y)$ , for some  $A$  and  $B$ . Substituting (b) into (a) we have  $y = \Sigma(A, \Sigma(B, y))$ , which by Leveling is  $\Sigma(A, B, y)$ . Substituting that into (b) yields  $x = \Sigma(B, \Sigma(A, B, y))$ , which by leveling, permutation, and absorption is  $\Sigma(A, B, y)$ , i.e.  $y$ .

20. What about unrestricted composition? Here we must distinguish between unrestricted composition in the  $\Sigma$  sense—i.e., the principle that  $\Sigma(x_1\dots)$  exists whenever  $x_1\dots$  do, and unrestricted composition in the sense of 'fusion', when 'fusion' is defined in terms of parthood (which in turn is defined in terms of  $\Sigma$ ). As for the former, Fine uses  $\Sigma$  as a term-forming operator and does not provide for empty terms in the background logic, and so simply presupposes that  $\Sigma$  composites exist; but of course, as he points out (note 11), one could relax this presupposition. As for the latter, even given unrestricted composition in the  $\Sigma$  sense, unrestricted composition in the Fu sense doesn't follow, as can be seen in the four-element model given in the text:  $a$  and

whose domain is the nonempty intervals of real numbers (closed or open on either side) drawn from  $[0,1]$ , plus nonempty unions of such intervals, and interpret  $\Sigma$  as union. Thus understood  $\Sigma$  is always defined, since a union of unions of intervals is itself a union of intervals. Absorption, collapse, leveling, and permutation all hold. Now consider:  $(0,1]$  is part of  $[0,1]$  in this model.  $((0,1] \cup [0,1] = [0,1]$ ; thus  $\Sigma((0,1],[0,1]) = [0,1]$ .) But  $(0,1] \neq [0,1]$ . Thus  $(0,1] \ll [0,1]$ . But any part of  $[0,1]$  overlaps  $(0,1]$ . (Let  $x$  be part of  $[0,1]$ , i.e.,  $[0,1] = x \cup \dots$ . Since  $x$  is in our domain, it contains some interval  $i$  from  $[0,1]$  as a (perhaps improper) subset. Divide  $i$  into thirds:  $i = i^l \cup i^m \cup i^r$ , so that  $i^m$  contains neither 0 nor 1.  $i^m$  is part of both  $x$  and  $(0,1]$ .)<sup>21</sup>

Does the fact that Fine’s principles don’t imply weak supplementation show that they do not *fully* capture “nothing over and above”? I don’t think so: weak supplementation bears an unclear relationship to that slogan. Furthermore, the only argument from the slogan to weak supplementation considered in the previous section ran through the uniqueness of fusions. But an operationalist could insist on interpreting the slogan only in terms of  $\Sigma$ . Thus understood it delivers only uniqueness of  $\Sigma$ -fusions, which does not imply weak supplementation (in the countermodel to weak supplementation, uniqueness of  $\Sigma$ -fusions is true). So far, then, the hypothesis that Fine’s principles fully capture the slogan remains standing.

However, the following model seems, at first glance anyway, to show that Fine’s principles do not, after all, fully articulate “nothing over and above”, and that they do not “constitute an analysis of the notion of mere sum”:



$$\Sigma(a, b) = \Sigma(b, c) = \Sigma(a, c) = \Sigma(a, b, c)$$

$b$  don’t have an Fu-fusion since the only object containing  $a$  and  $b$  as parts is  $d$ , but  $d$  has a part that doesn’t overlap either  $a$  or  $b$ , namely,  $c$  (compare Simons (1987, 32)). The counterexample also goes through under  $\text{Fu}^*$ , but unrestricted composition in the  $\text{Fu}^{\text{lub}}$  sense follows from Fine’s principles (and unrestricted  $\Sigma$ -composition) since  $\Sigma(x_1, \dots)$  counts as a least-upper-bound fusion of  $x_1, \dots$  (Obviously each  $x_i$  is part of  $\Sigma(x_1, \dots)$ ). Next suppose that each  $x_i$  is part of some  $y$ . Then by the definition of ‘part’ in terms of  $\Sigma$ ,  $(*) y = \Sigma(x_1, B_1) = \Sigma(x_2, B_2) = \dots$ . But since by absorption and collapse,  $y = \Sigma(y, y, \dots)$ ,  $(*)$  yields  $y = \Sigma(\Sigma(x_1, B_1), \Sigma(x_2, B_2), \dots)$ ; then by leveling and permutation,  $y = \Sigma(\Sigma(x_1, x_2, \dots), B_1, B_2, \dots)$  and so we have  $\Sigma(x_1, x_2, \dots) < y$ .

21. Note that in this model, both  $(0,1]$  and  $[0,1]$  are fusions in the ‘Fu’ sense of  $(0,0.5), [0.5,1]$ , so uniqueness fails.

Fine's four principles are satisfied in this model. But how can a single object,  $d$ , be nothing over and above, or a mere sum of,  $a$  and  $b$ , while also being nothing over and above, or a mere sum of,  $b$  and  $c$ ?

There is a complication, though. In a discussion of certain other composition operations (which obey principles other than summative identity), Fine says that "the only identities that should hold are the ones that can be shown to hold on the basis of the defining principles", and he continues in a footnote: "To put it algebraically, the intended model for the principles should be isomorphic to a 'word algebra' over the 'generators' or given elements" (575).<sup>22</sup> Now, Fine does not mention this further claim in his discussion of mere sums. Nevertheless, one possible view would be that the analysis of mere sum should include this further claim, and not merely the four principles. This would rule out my four-element model, since that model makes identifications (such as  $\Sigma(a, b) = \Sigma(b, c)$ ) that cannot be shown to hold on the basis of Fine's four principles (better: my model is not isomorphic to the appropriate word algebra over  $a$ ,  $b$ , and  $c$ ). Now, it is unclear how this view will work in general, since talk of generators will need to be explained in the case of gunk; but in any case it is certainly worthy of further exploration.

Let me close my discussion of Fine with a tangential point. Fine gives powerful arguments in favor of the operational approach, but here is a consideration that cuts the other way. In the operationalist approach to set theory (similar remarks apply in other cases), in which a set-builder operation  $\Sigma_{\text{set}}$  replaces  $\in$  as the primitive notion, logical ideology beyond that of first-order logic will be needed to formalize expressions " $\Sigma_{\text{set}}(x_1, x_2, \dots)$ " with infinitely many operands. Such resources are unneeded in familiar first-order set theories in which  $\in$  is the primitive notion. Of course, there are arguments—such as in Boolos (1984)—that first-order set theories are inferior to set theories formulated using stronger resources, such as plural quantification. But the operationalist's need for non-first-order resources is deeper: without them, the operationalist cannot even formulate the most basic claims about infinite sets. Moreover, the operationalist needs more powerful resources: the language must allow  $\Sigma$  expressions with lists of operands of arbitrarily high cardinality. If set theory is intended as a fundamental theory, the complexity of the associated logical ideology may well be significant.

---

22. The algebraic clarification of the view is needed; otherwise one might object: "let ' $e$ ' and ' $f$ ' name the same thing; ' $e = f$ ' is then true but does not follow from the defining principles".

### 3. *Composition as identity*

Let us turn next to a rather exciting—in both a good way and a bad way—attempt to articulate our slogan, namely, “composition as identity”. On this view, the intimate “nothing over and above” relation between a thing and its parts is just identity. Here are Megan Wallace and Donald Baxter in favor of this idea:

If the chair is distinct from the seat and the leg, then we are committed to co-located objects. The chair is a material object that occupies region,  $R$ . The seat and the leg are material objects that occupy region,  $R$ . This is ... complete spatial overlap: there is no place that the chair is that the seat and leg are not, and there is no place that the seat and the leg are that the chair is not. Since complete spatial co-location is unwelcome, then perhaps the seat and the leg are *not* distinct from the chair. (Wallace, 2011, 804)

Suppose a man owned some land which he divides into six parcels. Overcome with enthusiasm for the Non-Identity view he might try to perpetrate the following scam. He sells off the six parcels while retaining ownership of the whole. That way he gets some cash while hanging on to his land. Suppose the six buyers of the parcels argue that they jointly own the whole and the original owner now owns nothing. Their argument seems right. But it suggests that the whole was not a seventh thing. (Baxter, 1988a, 579)

The doctrine of composition as identity that I will consider does not say that for each part, the whole is identical to that part.<sup>23</sup> It says, rather, that the whole is identical to the parts *taken together*. To state the view more precisely, we employ irreducibly plural quantification (Bohlos, 1984). First redefine  $Fu$  in plural rather than set-theoretic terms:

$$x \text{ Fu} Y =_{\text{df}} \forall z (Yz \rightarrow z < x) \wedge \forall z (z < x \rightarrow \exists w (Yw \wedge Owz))$$

(“anything that is one of the  $Y$ s is part of  $x$ , and each part of  $x$  overlaps something that is one of the  $Y$ s”) and then state the doctrine as follows:

**Composition as identity** If  $x \text{ Fu} Y$  then  $x = Y$

On this view, the identity predicate, ‘=’—which expresses the one and only sort of identity—can take either plural or singular arguments on either

---

23. Baxter (1988a, b) himself defends a more radical version; see Turner (2014) for discussion.

side. Identity can hold one-many, and does so in the case of a composite and its many parts.

We began by noting the apparent incoherence of the slogan. We asked, rhetorically, how can I be nothing over and above my parts when they are many and I am one? Articulating the slogan as classical mereology avoided incoherence by reducing the slogan's role to a mere spirit or picture guiding the classical mereologist's choice of axioms; nothing directly corresponding to the slogan appears in the official theory. Similar remarks apply to Fine's theory. But composition as identity embraces the slogan more directly, by claiming that the whole is nothing over and above the parts in the perfectly literal sense of being literally identical to them. To those who feel the pull of composition as identity, this is the source of the attraction. As corollaries, Wallace's co-location may be avoided, and we may perhaps explain why Baxter's greedy owner now owns nothing.<sup>24</sup>

Defenders of composition as identity normally accept classical mereology. Composition as identity would be even more attractive if classical mereology *followed from* it—or at any rate, if all those parts of classical mereology that can be regarded as being underwritten by the slogan followed from it. For then, composition as identity could be regarded as *the* way to make the slogan precise. But whether this can be done depends on the underlying plural logic one accepts. In section 3.2 of Sider (2007) I used arguments like those in section 1 to derive all of classical mereology except unrestricted composition from composition as identity. But my arguments assumed a plural logic with a primitive plural-term-forming operator.<sup>25</sup> In a more typical logical setting, the distinctive plural logical ideology consists solely of plural quantifiers and variables, plus the predicate 'is one of'; and in this setting I see no way to reconstruct my arguments. Plural referring terms are usually eliminable using the prin-

---

24. In earlier work (2007) I complained that composition as identity cannot explain claims that are distinctive about parthood concerning the relation between a whole and a single one of its parts individually, such as certain "inheritance" principles. I considered the response that the inheritance principles are just analytic, and replied that even if they are, we need an explanation of why alternate notions are intuitively bizarre (e.g., a notion of 'location' on which one is "located" wherever any of one's relatives are located). But, it may be replied, all that's needed is a story about what is special about the meanings we do adopt, and this can be given by the defender of composition as identity: they're special in that they're defined in terms of parthood (rather than being a relative, for instance), and parthood is distinctive because of how it's connected to identity!

25. Furthermore, I did not consider the question of whether 'fusion' in composition as identity should be defined as  $Fu$ ,  $Fu^*$ , or  $Fu^{lub}$ .

ciple of plural comprehension, but full-strength comprehension is false given composition as identity. The natural weakening of comprehension is to “fusion-closed” pluralities (see Sider (2014)); but then certain principles of mereology must already be in place in order to apply this principle. (This problem does not confront the arguments of section 1, provided ‘nothing over and above’ is not claimed to obey an analog of Leibniz’s Law.) So perhaps the best composition as identity has to offer to the project of clarifying “nothing over and above” is an *addition* to classical mereology.

#### 4. *Composition as identity and unrestricted composition*

The relation between composition as identity and unrestricted composition is perplexing. Given my official formulation, composition as identity clearly does not imply unrestricted composition, not directly anyway. As officially formulated, composition as identity says merely that if there is a fusion of the *Xs*, then that thing is identical to those *Xs*; this apparently leaves open that the *Xs* simply have no fusion.<sup>26</sup> But there remains a persistent feeling that the intuitive idea of composition as identity really does imply unrestricted composition. If the whole is just its parts, then how could a whole fail to exist when the parts do? Perhaps the official formulation is too weak?

In section 1 I postponed discussion of whether unrestricted composition is underwritten by the slogan that “the whole is nothing over and above the parts”; now I can say why. On the one hand, the slogan seems to say merely that if there is a whole composed of certain parts then that whole stands in the “nothing over and above” relation to those parts, which apparently leaves open that there simply is no whole composed of those parts. On the other hand, *shouldn’t* the slogan be understood so as to imply unrestricted composition?

Speaking for myself, I find it hard to shake the feeling that unrestricted composition is part of the intuitive picture that gives rise to the slogan, and also part of the intuitive picture that makes composition as identity so alluring. Thus I am moved to inquire: is there any view in the vicinity of composition as identity—perhaps a strengthened form of composition as

---

26. I pointed this out in Sider (2007, 61), and then went on to mention (without endorsing) certain indirect routes from composition as identity to unrestricted composition.

identity, or perhaps a different view that is nevertheless in its spirit—from which unrestricted composition just falls out as a consequence?

#### 4.1 *Biconditional composition as identity*

Formulating composition as identity “biconditionally”, as the claim that  $x$  Fu $Y$  if and only if  $x = Y$ , doesn’t do the trick. Biconditional composition as identity doesn’t imply that some  $x$  fuses any given  $Y$ ; we need the added premise that for arbitrary  $Y$ , there is some  $x$  such that  $x = Y$ .

#### 4.2 *Identifying composition with identity*

Maybe composition as identity should be construed so as to identify, not composites with their parts, but rather the composition relation with the identity relation—with, that is, the plural identity relation which can relate pluralities.<sup>27</sup> Take any  $Y$ s. By the reflexivity of plural identity,  $Y = Y$ . But if the identity relation is the composition relation, then given suitable connecting premises one ought to be able to conclude that  $Y$  is composed of  $Y$ .

One problem with this view is that it seems obviously wrong. The composition relation is defined in terms of parthood; how can it also be the identity relation? It may be replied that relations are more coarsely individuated than we normally think; but I don’t quite see what intuitive picture this would come from.

A second problem is that the conclusion of the argument is merely that for any  $Y$ s, those  $Y$ s are composed of some  $Y$ s. But what we wanted is the conclusion that for any  $Y$ s, there is some  $x$  (singular) that is composed of the  $Y$ s. Identifying the composition and identity relations gets us nowhere.

#### 4.3 *No objective distinctions of number*

Here’s a (still) more bizarre attempt. Consider the following intuitive line of thought. For any  $Y$ s, there are some  $X$ s such that  $Y = X$ , namely the  $Y$ s themselves. So we have  $\forall Y \exists X Y = X$ . Now, the gap we’re struggling with is: how do we get from there to  $\forall Y \exists x Y = x$ ? (If we could, then given biconditional composition as identity we’d be home free.) Answer: there isn’t any objective distinction between manyness and oneness. The standard language of plural quantification makes a grammatical distinction

---

27. McDaniel (2014) mentions this view.

between singular variables  $x$  and plural variables  $Y$ , and thus is problematic: it makes a notational distinction where there is no corresponding distinction in reality. English too makes the unfortunate grammatical distinction between singular and plural. A better language just has one sort of variable:  $\alpha, \beta \dots$ . This language's " $\exists\alpha$ " can be read indifferently as "there is something,  $\alpha$ , such that ..." or "there are some things, the  $\alpha$ s, such that ...".

In the preferred language of this account, there's nothing like the gap we were discussing above. We have only  $\forall\alpha\exists\beta\alpha = \beta$ , which will be regarded as a logical truth about identity.

But there is a related gap. If the slogan requires unrestricted composition to be true, it surely also requires the following to be true:

**Existence of upper bounds**  $\forall\alpha\forall\beta\exists\gamma(\alpha < \gamma \wedge \beta < \gamma)$   
 "Any two 'things' are contained as parts by some further 'thing'"

What insures that this claim is true?

It may be responded that the theory is to include the axioms of classical mereology, including the principle of unrestricted composition, once those axioms are rewritten using the new variables. The existence of upper bounds follows from classical mereology thus understood.

At this point one wonders how the view differs from plain old classical mereology. But its defender might respond as follows. The comprehension schema for ordinary plural logic is the following, where  $\phi$  may be replaced by any formula containing just  $x$  free:

**Comprehension**  $\exists x\phi \rightarrow \exists Y\forall x(Yx \leftrightarrow \phi)$   
 "Provided there's at least one  $\phi$ , there are some things such that something is one of them iff it is a  $\phi$ "

Instances of comprehension are often regarded as having a certain distinctive status.<sup>28</sup> They are thought of as being easy to know, not in need of explanation, as being "trivial", as being logical truths, and so forth. If you are committed to the existence of at least one kangaroo, it is said, then accepting the existence of some things that are all and only the kangaroos is no further commitment. The defender of the view we are considering might insist that unrestricted composition—understood using the new number-indifferent variables—has that same status. The "objects" on her

---

28. See Lewis (1991, 3.2).

view, after all, are objectively no different from “pluralities”; and she might make the additional claim that the parthood relation is not objectively different from the is-one-of relation.

The idea, then, is to insist on the propriety of thinking of the variables *both* as plural (thus making unrestricted composition not a big deal) and as singular (thus allowing us to think of those variables as genuinely denoting composite objects).

Notice that composition-as-identity’s funny business with identity is no longer needed! We don’t have any plural/singular distinction anymore, so we don’t even have a way to say  $X = y$ .

My own view is that the alleged status of Comprehension is largely a mirage. For instance, even if accepting pluralities requires no further *ontological* commitment, it does require a further *ideological* commitment, which is a coequal way of “going out on a metaphysical limb” (Sider, 2011, 9.15).

But setting such concerns aside, the number-indifference view seems vulnerable to the following objection. Let  $x$  have no proper parts. Then, it would seem,  $x$  is objectively one. So the intuitive core of the view, that there is no objective difference between many and one, is untenable.

#### 4.4 *Fact identification*

Defenders of composition as identity sometimes say things like this: “you can think of the many as many, or as one”. But this actually suggests something quite different from a claim of many-one identity, namely that when we speak of many things existing, the fact that we are describing may also be described as a fact that involves the existence of one thing that is composed of those many. The nature of the fact itself does not call for a single fixed description.

On this view, the surprising identification is *not* between the one whole and the many parts; rather, the surprising identification is at the level of facts: the *fact* that the many exist is identical to the fact that there is also a one composed of them.<sup>29</sup> The view is not that things don’t have a fixed nature as being many or one; it’s that the facts don’t have a fixed nature as being facts about many or one.

The idea that facts have no fixed ontological form has tempted many. There is, for instance, the idea of “content recarving” in neoFregean phi-

---

29. I don’t, though, think of the view as requiring a robust commitment to facts.

losophy of mathematics. Agustín Rayo (2013) has recently articulated a version of this view; in his terms we could put the claim about parts and wholes as follows: for the parts to exist *just is* for the whole to exist in addition to those parts. There is also Eli Hirsch's (2011) quantifier variance, according to which there are multiple ways to understand quantifiers, each of which is adequate to describing the world. In Hirsch's terms, we could put the claim thus: no matter how we are initially using the quantifiers, if we can say truly that certain parts exist, then we are free to adopt a new, extended quantifier meaning under which we can say that there also exists a whole composed of those parts.<sup>30</sup>

The fact-identification idea could also be applied to various forms of *decomposition*. Just as a whole composed of given entities automatically exists, it might be said, the parts of a given whole under any chosen decomposition also automatically exist. For a fact involving a spatially extended object, say, can always be redescribed as a fact involving a left half and a right half.<sup>31</sup>

At the start of the paper the target sense of 'nothing over and above' was said to be a relation between a thing and its parts. Thus the fact-identification approach does not really capture that target sense. Nevertheless it strikes me as doing a lot of justice to some of the "nothing over and above" rhetoric. But those who find that rhetoric congenial may not feel the same way about content recarving or quantifier variance.

Some similar remarks apply to other "fact-level" interpretations of the slogan, such as the view that the fact of the whole's existence is grounded in the fact of the parts' existence, as defended by Ross Cameron (2014): though it does not really capture the target sense of 'nothing over and above', it does do justice to some of the rhetoric. Also in this vein, consider Trenton Merricks's (2001, 11f.) point that the elusive claim that a thing is nothing over and above its parts might simply be taken to say that the thing doesn't exist, only the parts do. Mereological nihilism—the view that composite objects do not exist—is therefore a possible interpretation of our slogan. And it becomes like the other fact-level interpretations we have been considering if the nihilist is willing to say: "although 'composites

---

30. Other views that are more or less in this spirit include those of Carnap (1950), Fine (2006, 2007), Jubien (1993, 2009), Thomasson (2007), and McDaniel (2013), who discusses the view that some entities are "no additions to being" in the sense of being in the range only of nonfundamental quantifiers.

31. I'm tempted to read Hofweber and Velleman (2010) as implicitly presupposing some such view.

do not exist' is true in the language I am now speaking, we can introduce a new language in which 'there exists something composed of the *Y*s' has the same truth condition as 'the *Y*s exist' has in the language I am now speaking, and thus is true" (Sider, 2013a, section 3). For in so saying, the nihilist has accepted a sort of quantifier variance, albeit without the usual claim that all the quantifier meanings are on a par, and arguably has also accepted a form of Cameron's grounding view.<sup>32</sup>

### 5. *Classical mereology again*

In section 4 we sought an interpretation of composition as identity, or of the slogan that the whole is nothing over and above the parts, that implies unrestricted composition. But perhaps what we sought was right under our noses all along. Classical mereology, our first interpretation of the slogan, implies unrestricted composition, because unrestricted composition is one of its axioms.<sup>33</sup>

This might seem too thin. The slogan, it might be thought, calls for composites to "automatically" exist whenever the parts do. Classical mereology does not deliver the automatic existence of composites; it merely asserts that they exist. But it's worth putting pressure on this notion of automaticity. What does it mean to say that something is automatically true, given certain other truths?

One sort of automaticity is material: universal truth-preservation, regardless of time or place. Since classical mereology says that *all* collections have fusions, no matter how they're arranged, the existence of a fusion in any given case *is* automatic in the material sense.

A stronger sort of automaticity is modal: truth-preservation regardless of the possible circumstances—necessary truth-preservation. But many defenders of classical mereology would say that its axioms aren't merely contingently true; they are necessarily true.<sup>34</sup> If so, then the existence of composites again counts as automatic. This may still

---

32. The latter subsumption is most apt under a "deflationary" approach to ground (in the sense of von Solodkoff and Woodward (2013)), of which I count my "metaphysical semantics" approach as an instance (Sider 2011, section 7.9, 2013b).

33. Unlike the other axioms, unrestricted composition was not "derived" in section 1 from the slogan, so it may be objected that only the other axioms constitute an interpretation of the slogan. But the persistent feeling driving the previous section suggests otherwise.

34. Cameron (2007) disagrees. My own view is that the necessity of unrestricted composition would amount to little more than its truth (Sider, 2011, chapter 12).

seem too thin. But on closer inspection, it's unclear what more we might want.

A putatively stronger sort of automaticity is logical: truth-preservation by virtue of logic—logical consequence. A disjunction is automatically true in this sense, given a true disjunct. Since the axiom of unrestricted composition in classical mereology is not a logical truth, it may be argued, we have identified an important sense of automaticity in which classical mereology does not deliver the automatic existence of composite objects. But whether this is correct depends on the nature of logical consequence.

According to Quine's definition, for example, one sentence logically implies another just when it's not the case that the first is true and the second is false, and moreover, that this remains so no matter how one uniformly substitutes expressions for nonlogical expressions in the two sentences.<sup>35</sup> So in essence, disjunctions are logically implied by their disjuncts because i) 'or' is a logical expression and ii) in fact, as it happens, disjunctions always are true whenever they have a true disjunct. Now, the thesis of unrestricted composition contains (when formulated in primitive terms) only standard logical expressions—quantifiers, the identity sign, and boolean connectives—plus the predicate  $<$  for parthood. So if  $<$  counts as a logical expression too, then every expression in unrestricted composition is a logical expression, in which case its mere truth would suffice for its logical truth, given Quine's view. Thus the only thing standing in the way of saying that the existence of the whole is *logically* implied by the existence of the parts is the claim that 'part' is not a logical expression—a claim that is questionable, and moreover of questionable depth.

Quine's conception of logical consequence is often regarded as being too weak, since it allegedly leaves out the modal element of logical consequence; but certain natural strengthenings don't significantly alter the picture. Consider, for instance, the view that  $p$  logically implies  $q$  iff: it's *impossible* for  $p$  to be true while  $q$  is false, and moreover, this remains so for all uniform substitutions for nonlogical expressions in  $p$  and  $q$ . This too does not significantly alter the picture, if classical mereology is necessarily true. For then the implication of the existence of the whole from the existence of the parts is either logical, or else just like a logical implication except for the—shallow, I think—fact that  $<$  does not count as a logical constant.

---

35. Quine (1960); p. 103 in Quine (1966).

We are seeking a sense of automaticity under which the slogan demands, but classical mereology cannot deliver, the automatic existence of composite objects. The two conceptions of logical consequence we have considered so far, which did not lead to the desired sense of automaticity, were reductive. So perhaps a primitivist conception of logical consequence would fare better. Perhaps what the slogan really demands, and what classical mereology cannot deliver, is that the existence of the parts logically imply, in the primitive sense, the existence of the whole. Relatedly, a follower of Fine (2001, 2012, 1994) might try to capture the target sense of automaticity by saying that what the slogan demands, and what classical mereology cannot deliver, is that the existence of the whole be *grounded* in the existence of the parts, in some primitive sense of grounding, or that it be of the *essence* of mereological composition (or quantification, or identity), in some primitive sense of essence, that unrestricted composition hold.

But it is doubtful that these proposals really scratch the itch we are feeling. To bring this out, imagine someone who claimed that the existence of the whole *is* primitively logically implied by, or essentially implied by, or grounded in, the existence of the parts. Intuitively, such primitivist claims would not deliver the sort of automaticity we want (or think we want). For it would surely be natural to demand of the primitivist: *how* does the connection hold? What is the *mechanism* by which the connection holds; what is it, specifically, about composition, or identity, or quantification, that is responsible for the connection holding, and how does this specific feature of composition, identity, or quantification bring about the connection? The bare claim that the connection holds, absent a mechanism, feels no more satisfying than classical mereology's bare assertion that unrestricted composition is true. What we were doing in section 4 was seeking such mechanisms; the primitivist claims we are now considering amount to saying that some such mechanism exists without identifying it.

But if the mechanism is what we're really after, to scratch the itch, then the account of the automaticity given by plain old classical mereology doesn't look so thin after all. Each view considered in section 4 attempted to identify some claim (about composition, identity, or quantification) that would imply unrestricted composition. But unrestricted composition itself is a claim that implies unrestricted composition. If we could derive it from some other fact about parthood, composition, identity, or quantification, that would be nice; but then that further fact itself wouldn't

have any deeper explanation. We can't escape unexplained explainers; at most, we can embed unrestricted composition in a unified and satisfying theory. The views discussed in section 4 attempt to do this, and perhaps they do it better than classical mereology. But classical mereology is also unified and satisfying in its own way.

#### REFERENCES

- Baxter, Donald 1988a: "Identity in the Loose and Popular Sense." *Mind* 97, 575–82.
- 1988b: "Many-One Identity." *Philosophical Papers* 17, 193–216.
- Baxter, Donald & Cotnoir, Aaron (eds.) 2014: *Composition as Identity*. Oxford: Oxford University Press.
- Bennett, Karen 2011: "Construction Area: No Hard Hat Required." *Philosophical Studies* 154, 79–104.
- 2015: *Making Things Up*. Oxford: Oxford University Press. Forthcoming.
- Boolos, George 1971: "The Iterative Conception of Set." *Journal of Philosophy* 68, 215–31. Reprinted in: Boolos 1998, 13–29.
- 1984: "To Be Is to Be the Value of a Variable (or to Be Some Values of Some Variables)." *Journal of Philosophy* 81, 430–49. Reprinted in Boolos 1998, 54–72.
- 1998: *Logic, Logic, and Logic*. Cambridge, Mass.: Harvard University Press.
- Cameron, Ross P. 2007: "The Contingency of Composition." *Philosophical Studies* 136, 99–121.
- 2014: "Parts Generate the Whole, but They Are Not Identical to It." In: Donald Baxter and Aaron Cotnoir, *Composition as Identity*. Oxford: Oxford University Press, 90–107.
- Carnap, Rudolf 1950: "Empiricism, Semantics and Ontology." *Revue Internationale de Philosophie* 4, 20–40. Reprinted in: *Meaning and Necessity: A Study in Semantics and Modal Logic*, 2nd edn. Chicago: University of Chicago Press.
- Fine, Kit 1994: "Essence and Modality." In: James Tomberlin (ed.), *Philosophical Perspectives 8: Logic and Language*. Atascadero, CA: Ridgeview, 1–16.
- 2001: "The Question of Realism." *Philosophers' Imprint* 1, 1–30.
- 2006: "Our Knowledge of Mathematical Objects." In: Tamar Szabó Gendler & John Hawthorne (eds.), *Oxford Studies in Epistemology*, volume 1. Oxford: Clarendon Press., 89–109.

- 2007: “Relatively Unrestricted Quantification.” In: Agustín Rayo & Gabriel Uzquiano (eds.), *Absolute Generality*. Oxford: Oxford University Press, 20–44.
- 2010: “Towards a Theory of Part.” *Journal of Philosophy* 107, 559–89.
- 2012: “Guide to Ground.” In: Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press, 37–80.
- Hirsch, Eli 2011: *Quantifier Variance and Realism: Essays in Metaontology*. New York: Oxford University Press.
- Hofweber, Thomas & Velleman J. David 2010: “How to Endure.” *Philosophical Quarterly* 61, 37–57.
- Hovda, Paul 2009: “What Is Classical Mereology?”. *Journal of Philosophical Logic* 38(1), 55–82.
- Jubien, Michael 1993: *Ontology, Modality, and the Fallacy of Reference*. Cambridge: Cambridge University Press.
- 2009: *Possibility*. Oxford: Oxford University Press.
- Lewis, David 1991: *Parts of Classes*. Oxford: Blackwell.
- McDaniel, Kris 2004: “Modal Realism with Overlap.” *Australasian Journal of Philosophy* 82, 137–52.
- 2009: “Structure-Making.” *Australasian Journal of Philosophy* 87, 251–274.
- 2013: “Degrees of Being.” *Philosophers’ Imprint* 13(19).
- 2014: “Compositional Pluralism and Composition as Identity.” In: Donald Baxter and Aaron Cotnoir, *Composition as Identity*. Oxford: Oxford University Press, 130–142.
- Merricks, Trenton 2001: *Objects and Persons*. Oxford: Clarendon Press.
- Quine, Willard V. O. 1960: “Carnap and Logical Truth.” *Synthese* 12, 350–74.
- 1966: *The Ways of Paradox*. New York: Random House.
- Rayo, Agustín 2013: *The Construction of Logical Space*. Oxford: Oxford University Press.
- Sider, Theodore 2007: “Parthood.” *Philosophical Review* 116, 51–91.
- 2011: *Writing the Book of the World*. Oxford: Clarendon Press.
- 2013a: “Against Parthood.” In Karen Bennett & Dean W. Zimmerman (eds.), *Oxford Studies in Metaphysics*, volume 8. Oxford: Oxford University Press, 37–93.
- 2013b: “Reply to Jonathan Schaffer.” *Analysis* 73, 760–70.
- 2014: “Consequences of Collapse.” In: Donald Baxter and Aaron Cotnoir, *Composition as Identity*. Oxford: Oxford University Press, 211–21.
- Simons, Peter 1987: *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- Thomasson, Amie L. 2007: *Ordinary Objects*. New York: Oxford University Press.

- Turner, Jason 2014: "Donald Baxter's Composition as Identity." In: Donald Baxter and Aaron Cotnoir, *Composition as Identity*. Oxford: Oxford University Press, 225–43.
- von Solodkoff, Tatjana & Woodward Richard 2013: "Noneism, Ontology, and Fundamentality." *Philosophy and Phenomenological Research* 87(3), 558–83.
- Wallace, Megan 2011: "Composition as Identity: Part 1." *Philosophy Compass* 6, 804–16.

ON BASIC MODES OF BEING:  
METAMETAPHYSICAL REFLECTIONS IN LIGHT OF  
WHITEHEAD, HUSSERL, INGARDEN, HINTIKKA

David WOODRUFF SMITH  
University of California, Irvine

*Summary*

Here I should like to explore a conception of ontological structure that has but few precedents. The idea is that certain *structures* are *basic* in the formation of the world, deep structures that define basic *modes of being* for things in the world, modes that make a thing the thing it is. We might think of this deep structure of a thing as an algorithmic formation of the thing itself.

I'll develop a variation on Whitehead's model of "concrecence", or "becoming" a concrete entity, drawing on Husserl's conceptions of dependence, dependent parts, and manifolds, and Ingarden's conception of modes of being. In the end I'll apply this model of concrecence to Husserl's analysis of the deep structure of consciousness, including time-consciousness and body-consciousness. Into this model I'll draw Hintikka's conception of intentional modalities, including perceptual possibility, to bring out a structure of modes-of-being for a conscious visual experience.

1.

Here I should like to explore a conception of ontological structure that has but few precedents. The idea is that certain *structures* are *basic* in the formation of the world, structures that define basic *modes of being* for things in the world, modes that make a thing the thing it is. We might think of this deep structure of a thing as an algorithmic formation of the thing itself, but we'll need to work up to that image.

We are used to the notion of deep structure in theories of language—from Frege and Quine in formal logic to Chomsky in formal grammars of everyday language. Here we seek a picture of deep structure in ontology, where familiar structures of everyday objects or situations or events are

founded on more basic structures that underlie the everyday, structures that define as it were the very constitution of everyday phenomena and their ways of being in the world.

A long-running tradition since Aristotle focuses on concrete individuals, or “substances”, and their properties—kinds, qualities, and relations. In that metaphysic the world is an array of concrete particulars bearing various properties that inform their matter. Concrete individuals occur in spacetime and their properties may change over time while the individuals themselves endure through time. A more recent tradition, in light of modern physics, focuses on events as more basic than physical individuals in nature. In this lineage Alfred North Whitehead sees the world as a flux of concrete events, or “actual occasions”, which are more fundamental than concrete Aristotelian “substances”. But there is another dimension of ontology, a deeper dimension, discernible in Whitehead’s metaphysic. And that dimension of “depth” is what I want to explore here.

Timely, as it happens, is an “untimely review” of Whitehead’s *Process and Reality* (1978/1929) published recently by Peter Simons (Simons 2013). That review provides just the actual occasion here to revisit Whiteheadian ontology. I concur with Simons’ appraisal that *Process and Reality* “has good claims to be the greatest metaphysical treatise for many years, certainly since the beginning of the twentieth century.” And accordingly we should plumb the depths of Whitehead’s visionary system for a glimpse of the deep structure of the world. Into this venture I’ll draw in Husserl’s conception of structures called manifolds and their roles in wholes bound together by ties of ontological dependence. From there I’ll draw further on Ingarden’s conception of modes of being (featuring dependence). Professor Simons has written extensively on the ontology of dependent parts, extending Husserlian themes (see Simons 1987, Simons 1992, and his contributions in Smith (ed.) 1982). And today themes of dependence and “grounding” are re-appearing in a different philosophical milieu. (See Correia and Schnieder (eds.) 2012, especially the essays by Kit Fine and Kathrin Koslicki. However, a particularly rich ontology of dependence is developed by Roman Ingarden in Ingarden 2013/1947, amplifying Husserl’s theory of dependent parts or “moments” with a theory of bona fide “modes of being”.)

In what follows I’ll develop a variation on Whitehead’s model of “concrecence”, or “becoming” a concrete entity, explicating the notion partly by adapting Husserl’s conceptions of dependence, dependent parts, and manifolds, and Ingarden’s conception of modes of being. In the end I’ll apply

this model of concrescence to Husserl's analysis of the deep structure of consciousness, including time-consciousness and body-consciousness. Into this model I'll draw Jaakko Hintikka's conception of intentional modalities (following Hintikka 1969), including perceptual possibility, to bring out a structure of modes-of-being for visual experience and its intentionality.

I'll be working from specific ideas drawn from Whitehead, Husserl, Ingarden, and Hintikka. For the most part, I'll simply assume reconstructions of those ideas in prior works, as my aims here are not interpretive exegesis but conceptual exploration. We'll be exploring how we might envision basic ontological structures, in ways that doubtless go beyond the writings of those authors that inspire this study. (For my reconstruction of Whiteheadian themes, relevant here, see Smith 2004, Chapter 7 on Whitehead's metaphysics and Chapter 8 on the metaphysics of basic categories. For my reconstruction of Husserlian theories, relevant here, see Smith 2013. On Hintikka models apropos of Husserlian intentionality, see Smith and McIntyre 1982. Here I'll sketch the relevant ideas as we go, but I'll not cite chapter and verse from either Whitehead or Husserl or Ingarden or Hintikka. Very broadly, the present essay looks to the foundations of ontology rather in the spirit of "metametaphysics": compare Chalmers, Manley, and Wasserman (eds.) 2009.)

## 2.

Let us set the scene by contrasting some very basic approaches to the fundamental structure of the world.

Aristotle held, in the *Categories* (Aristotle 1995), that the most fundamental type of things ("not in combination") are concrete individuals called "[primary] substances" (*ousia*, beings). Of substances several types of things are "predicated": their species (or "secondary substances"), their quantities (distinguishing them in number); their qualities, their relations (or "relative" properties), their temporal locations, their spatial locations, and so on. ... The world, if you will, comprises or is formed from all the substances, inhering in which are "predicates" of these varied types. A nominalist or particularist variation on this metaphysic either dismisses the "predicates" (universals) or compresses them into their particularized instances within substances.

Against this long-running traditional ontology, Ludwig Wittgenstein held, in the *Tractatus* (Wittgenstein 1974/1921), that the world is all the

“facts”, defined as existing states of affairs or situations. Not objects!—but states of affairs. Objects (individuals) stand in relations (or have monadic properties) and thus form states of affairs (*Sachverhalten*). A state of affairs is a combination of objects, wherein the objects are appropriately related. Such states of affairs, if existing, form the world. But the world consists, Wittgenstein held, not in an array of objects (individuals, particulars), but in an array of all the existing atomic states of affairs. ... While we may talk of conjunctions or disjunctions or pluralities of these states of affairs, only the atomic ones exist. Our logically complex propositions (“A and B”, “A or B”, “not A”, “if A then B”) are made true by the relevant atomic states of affairs, without need for more complex states of affairs.

Against both the Aristotelian model and the Tractarian model, Whitehead held, in *Process and Reality* (Whitehead 1978/1929), that the world consists of all the concrete events, or “actual occasions”. Not all the substances!—substantial concrete individuals (with inhering properties or “predicates” of several basic types). Nor all the atomic states of affairs!—formed from individuals and their properties or relations. Rather, Whitehead pressed, the world—embracing spatiotemporal reality—is formed fundamentally from all the actual occasions, or concrete events. Not from all the substances, and not from all the atomic states of affairs involving substances. ... By the way, it was Bertrand Russell along with Wittgenstein who articulated the metaphysic of “facts” or states of affairs, Russell calling his version “logical atomism”. Russell and Whitehead had written *Principia Mathematica* (1910–1913); subsequently Russell had met with Wittgenstein; then Russell and Wittgenstein formulated their respective versions of an ontology of facts or states of affairs. When Whitehead turned from mathematics to metaphysics and produced *Process and Reality*, he was seeking deeper structures of reality than either objects or states of affairs. Or so I shall assume.

### 3.

Now, the “processes” involved in Whitehead’s model of “process and reality” are not what we normally call “processes”, that is, temporal processes wherein one event follows another in time, following in some orderly pattern (a volcano erupting, a carpenter building a cabinet). Rather, for Whitehead, events in spacetime—actual occasions in the “extensive continuum” of spacetime—*come into being*, or *become*. The *process* by which a

concrete event *becomes* that particular event is not a temporal process, but as it were a *mathematical process* defining the *structure* of that event's place in a nexus of related events. An event is the event it is only *in relation to* its kindred events: quite as relativity physics defines a location in spacetime in terms of its relation to other locations in spacetime, whence antecedent events are in the "light cone" of a current event, taking physical events as swirls of matter-energy in spacetime.

But while relational models of space and time are familiar, Whitehead's model of the process of "becoming"—an event's coming-into-being *per se*, whereby it takes its place in spacetime—is something else: a deeper structure of reality.

To gloss Whitehead's doctrine: "Becoming is not temporal; temporality becomes." In Whitehead's own words:

In every act of becoming there is the becoming of something with temporal extension; but ... the act itself is not [temporally] extensive, in the sense that it is divisible into earlier and later acts of becoming which correspond to the [temporally] extensive divisibility of what has become. (Whitehead 1978/1929, 69)

Whitehead's "becoming" is not a temporal becoming, but rather an ontological becoming, an *ontogenesis*, an emergence into being. Whitehead called his overall metaphysic the philosophy of "organism", treating the universe itself as an organism. Of course, biology famously distinguished between the phylogenesis and the ontogenesis of an organism: where ontogeny recapitulates phylogeny. However, the ontogenesis we are exploring is a structure that emerges, as it were, mathematically.

Metaphorically, we may say, on the Whiteheadian model: *temporal becoming* is a "horizontal" structure, stretching over the flow of things in time or (better) in spacetime, whereas *ontological becoming* is a "vertical" structure, stretching from the deep structure of things upward into their being in the spacetime continuum.

To amplify: *Reality* is the spatiotemporal array of concrete events in relation to kindred concrete events. And the process of ontological *becoming* is the pattern or *structure* wherein a concrete event is defined as being formed in relation to its kindred concrete events within the spatiotemporal spread of *reality*. ... We should honor the economy of Whitehead's formulation above. For his full elaborations are, shall we say, most intricate. We'll need to sketch something of Whitehead's picture of how an event is linked to its kin in a "nexus" of relationships. But before we get lost

in the trees I want to pull back for a satellite view of the forest itself. In that view we see what “becoming” an “actual occasion” amounts to in the Whiteheadian scheme.

#### 4.

In *Process and Reality* Whitehead posits different types of ontological categories (Whitehead 1978/1929, 20ff). I want to draw upon only certain elements in his scheme, simplifying the Whiteheadian vision for present purposes.

The first type of category Whitehead calls “the Category of the Ultimate”. This category is the fundamental way in which a “creative advance” leads from “many” entities to “one” “novel” entity. In this process the one entity “becomes”: it *becomes* that one entity as its many antecedent entities come “together” in a process of “concrecence”. This *concrecence* is the fundamental manner of formation of any entity, the “production of novel togetherness” (21), as the entity comes together in its being. *Becoming*, becoming an entity through *concrecence*, is the most fundamental—the ultimate—mode of being of any entity whatsoever. And so, Whitehead holds, “This Category of the Ultimate replaces Aristotle’s category of ‘primary substance.’” (21)

The second type of category Whitehead calls “The Categories of Existence”. These categories include: “Actual Entities”, or “Actual Occasions”; “Prehensions”, or “Concrete Facts of Relatedness”; and “Eternal Objects”, or “Forms of Definiteness”. Actual occasions are concrete events, eternal objects are like Platonic forms, and prehensions are the types of relation that an actual occasion has, on the one hand, to other actual occasions on which it depends and, on the other hand, to eternal forms on which it depends. (Whitehead includes some other categories of existence, but these three categories of existence, or existents, are the only ones I’ll consider here. Whitehead also posits two other types of categories, which I’ll ignore for our purposes.)

The principal notion I want to draw from Whitehead’s system is the notion of becoming an actual entity through concrecence (or derivation) from other actual entities and from appropriate forms. This pattern of becoming we should see as an abstract or formal ontological structure, a structure that applies to actual occasions and in its way brings them into existence or actuality.

## 5.

For Whitehead, then, the ontological process of *becoming*—becoming a particular concrete entity—follows a pattern of formation Whitehead calls “conrescence”. I’ll generally prefer the latter term, as does Simons (in Simons 2013), precisely because Whitehead’s term “process” has often been misconstrued as meaning a temporal process. Whence “process ontology” is then taken to focus on events and their temporal flux, drawing on Heraclitus’ ancient vision. By contrast, I think what is truly original and truly “deep”—if the metaphysical Zeitgeist of today allows us to look once again into the “depths” of being—is the fundamental nature of conrescence.

*Conrescence*, I should like to propose, can be seen, in the Whiteheadian model, as an *algorithmic process*. In other words, think of a mathematical process: say, the process of evaluating  $n!$ —whatever number (whole integer)  $n$  is, multiply  $n$  times  $n-1$ , then multiply that number by  $n-2$ , and so on, until you are multiplying by 1, and then you have the value of  $n!$ , at the end of this process of evaluation. Abstract the mathematical form of calculation: at issue is not the human mental activity of calculating, but the ideal form of transformation from one value to another, the form dubbed an algorithm. Now think of the mathematical process by which an array of events is transformed, algorithmically, into a particular actual event  $E^*$ . The *structure* of the relationship between  $E^*$  and its antecedent events—those “before”  $E^*$  in the mathematical transformation—is the form of the conrescence of  $E^*$ .

Think in terms of a diagram. Take a particular concrete event  $E^*$  in spacetime. Consider the antecedents  $\langle E_1, \dots, E_n \rangle$  of  $E^*$ . Then the conrescence of  $E^*$  is a transformation from  $E^*$ ’s antecedents to  $E^*$  itself:

$$\langle E_1, \dots, E_n \rangle \text{ —C—} \rightarrow E^*$$

The transformation  $C$  carries the array of antecedents of  $E^*$  into  $E^*$  itself. This mathematical transformation is the *conrescence* of  $E^*$ , its *becoming*  $E^*$  itself. And that process of becoming  $E^*$  is not itself temporal. Rather, the conrescence of  $E^*$  places  $E^*$  in the spatiotemporal extension of the world—in its spatiotemporal relations to its antecedents. ... This sketch is of course a simplistic idealization, as the Whiteheadian model will ultimately have to confront issues of continuity and of quantum states of a system, so that the antecedents of  $E^*$  in reality will be as complex as are specified by mathematical physics in the long run.

Then, we may say, the existence of  $E^*$  is *ontologically dependent* on the *concrecence structure*:

$$\langle E_1, \dots, E_n \rangle \text{ ---C---} \rightarrow E^*.$$

If you will, that structure is the deep structure of the event  $E^*$ , without which  $E^*$  could not be and could not be:  $E^*$ . According to the Whiteheadian model.

In order better to bring out this deep structure, however, I'll co-opt and adapt some ideas from Husserl's ontology. (The notion of mathematical structures is central to Charles Parsons' work in philosophy of mathematics; see Parsons 2012 on historical themes leading from Kant to Husserl. The very idea of mathematical structures realized in the world is explored in Smith 2002, reflecting on Husserl's conception of "manifolds".)

## 6.

Many years before Whitehead's *Process and Reality* (1929), Edmund Husserl—another mathematician turned philosopher—laid out a complex categorial ontology. In his *Logical Investigations* (Husserl 2001/1900), Husserl developed a theory of parts and wholes featuring a key role for ontological dependence. And in his *Ideas I* (Husserl 1983/1913), he recast his ontology with further distinctions. (I rely here on details in my own reconstruction of Husserl's categorial ontology in my *Husserl* 2013.)

For Husserl, a concrete individual is typically a whole with two types of parts: *independent* parts, which can exist independently of the whole, like the wheels of a bicycle; and *dependent* parts, or "moments", which cannot exist apart from the whole, like the particular instance of red in a red bicycle. This notion of *moments* follows Aristotle's notion of "qualities" that cannot exist apart from a substance, like the particularized red in a particular red object. For Husserl, the ideal essence Red is a nonspatiotemporal entity, shareable by different red objects, while a particular instance of Red in a particular red object is a moment of the object, a part that can exist only in and where and when the object exists. So Husserl would have his Plato and his Aristotle too. (Property-instances are often called "tropes" these days.)

Now, Husserl amplifies his theory of universals and their instantiation with a theory of *ontological dependence*. A *moment*  $M$  of an individual  $X$  is

ontologically dependent on X, in that: the essence of X is such that M could not exist unless X exists. And a property-instance in a concrete individual X is a moment of X. Further, the essence of X is such that X could not exist unless M exists. The concrete individual X is thus a whole that cannot exist unless said moments of X exist and, vice versa, said moments of X cannot exist unless X exists. ... Husserl defines a “precise [*prägnant*] whole” as a whole wherein: each moment can exist only if each other moment exists, and the whole can exist only if each of the moments exist, and of course each moment can exist only if the whole exists—dependence tying all the parts and whole together.

Along these lines, Husserl would hold that a concrete object, say, a red ball of a certain diameter, cannot be that very object as it is now unless it and its property-instances are as they are now. Husserl’s paradigm of moments is, famously, that of particularized universals that are dependent parts of a concrete individual. However, other applications of the theory of moments are feasible. We might take a social community to be a whole comprising persons that are moments of the community. Thus, I am a Professor in the University of California, and UC is a community in which I serve as a dependent part of UC: I instantiate an instance of the type Professor, and that Professor-instance is realized as a moment of UC.

This Husserlian model of moments can be seen as an elaboration of a broadly Aristotelian ontology of “substances” conceived now as wholes formed from particularized universals.

We could then define the *concrescence* of a concrete individual à la Husserl as the way in which the *whole* individual is formed from its *moments* that are instances of properties in the essence of that individual (properties including kinds, qualities, and relations). That is, the concrescence of an individual X would be the mathematical transformation:

$$\langle p_1, \dots, p_n \rangle \text{ ---C---} \rangle X,$$

where  $p_1$  is an instance of property  $P_1$ , ... ,  $p_n$  is an instance of property  $P_n$ . Then, we might say, the individual X is *ontologically dependent*—or à la Husserl *founded*—on the concrescence structure:

$$\langle p_1, \dots, p_n \rangle \text{ ---C---} \rangle X.$$

In other words, we can use the Whiteheadian notion of concrescence to articulate the Husserlian notion of how a concrete individual is *formed*

from its concrete moments. Indeed, that structure itself can be modeled on the Husserlian notion of a “manifold”, a *structured* whole. For Husserl, the individual X is formed as a whole comprising its property-instances  $\langle p_1, \dots, p_n \rangle$ . But here we consider a more complex whole wherein concrescence carries  $\langle p_1, \dots, p_n \rangle$  into X itself: the relevant whole being, not simply the individual X with dependent parts  $p_1, \dots, p_n$ , but the structured whole wherein X is linked by dependence to  $p_1, \dots, p_n$ . If you will, the concrescence structure of X is the “manifold” structure:  $\langle X, \langle p_1, \dots, p_n \rangle \rangle$  where X is ontologically dependent on  $\langle p_i \rangle$ . Within that structure concrescence carries those property-instances into the concrete individual that bears them as moments.

So an individual’s concrescence structure defines the ways in which the individual *becomes* that very individual—as it were, in logical space or (better) in ontological space. Those ways are its *modes of being* thus-and-so: whence its ways of becoming that very individual—not a temporal process of coming into being (say, as a red ball is produced in a factory), but a formal process defining what the object is, what it is as it is here and now.

The structure of concrescence would be a part of the formal ontology of an individual. Now, Husserl posited formal categories such as Individual, Property, State of Affairs, Number, Manifold, etc. But Husserl applied these *formal* categories to three *material* categories called “regions”: Nature, Consciousness, and Culture (or “Spirit”, *Geist*). Things in nature are characterized by spatiotemporal properties, whereas acts of consciousness are characterized by intentionality (being directed toward things), and activities of persons in cultural communities are characterized by social or intersubjective relations. Then, following Husserl configured now along Whiteheadian lines, we would say that patterns of concrescence define different *modes of being* for things-in-nature, for acts-of-consciousness, and for social activities-in-culture: drawing respectively on instances of properties appropriate to nature, consciousness, and culture.

In this broadly Husserlian model, we might say an individual is *constituted* by its property-instances, that is, in the underlying concrescence structure of the individual. However, Husserl himself used the term “constitution” for the pattern of “noematic” meanings entertained in acts of consciousness, meanings that “constitute” the object *as intended* in consciousness, a pattern of *meanings* that aim at (without reducing to) this object’s pattern of concrescence in the world. We’ll return to the Husserlian

form of constitution below, reconfiguring phenomenological constitution in terms of concrescence.

7.

Now let us return to Whitehead's own model of an actual occasion and its concrescence.

Whitehead introduced a technical term for the fundamental linkage between a concrete event, or actual occasion, and its antecedent entities. He called this linkage "prehension": alliterating with "apprehension" or "comprehension", but meaning a "prehensile" grasp, as a monkey's prehensile tail grasps something. Each (currently) actual occasion *prehends* or "feels" a range of other (formerly) actual occasions and also prehends or "feels" a range of "eternal forms" or properties. So, for Whitehead, the *concrescence* of a concrete event  $E^*$ — $E^*$ 's *becoming*  $E^*$ —consists in the transformation from a range of prehended concrete events  $\langle E_1, \dots, E_n \rangle$  plus a range of prehended forms or properties  $\langle P_1, \dots, P_m \rangle$ . In the style we are using, the concrescence of  $E^*$  takes the form:

$$\langle \langle E_1, \dots, E_n \rangle, \langle P_1, \dots, P_m \rangle \rangle \text{ ---C---} E^*.$$

The properties  $P_i$  characterize  $E^*$  in a Platonistic mode-of-being, while the events  $E_i$  characterize  $E^*$  in a Heraclitean mode-of-being. In fact, Whitehead alludes to Plato's *Timaeus*, which perhaps blends the eternal forms into the "process" of becoming as  $E^*$  prehends its antecedent events  $E_i$ .

Remember that this process of concrescence, the becoming of  $E^*$ , is not itself temporal or spatiotemporal. Rather, in virtue of its concrescence,  $E^*$  takes its place in spacetime. And its ontological antecedents  $E_i$  take their places in spacetime temporally-prior to  $E^*$ . Within its concrescence,  $E^*$  also prehends its eternal forms or properties.

Interestingly, Whitehead holds that each actual occasion  $E^*$  together with its prehended occasions  $E_i$  form the *actual world* of  $E^*$ . Only  $E^*$  itself is *actual*, that is, actualizing *now* in spacetime in relations of prehension to *now-past* events  $E_i$ . The structure of relativity is thus assumed in this Whiteheadian picture of concrescence. ... Moreover, Whitehead's model of the actualizing of an actual occasion, producing the event's *now* being actual, arguably entails a form of the doctrine called "presentism", whereby only the present moment or occasion is actual. (A recent model of tem-

porality and becoming distinguishes temporal, absolute, and atemporal types of becoming: see the discussion in Fiocco 2014, citing C. D. Broad's notion of absolute becoming, all of which I believe echoes Whiteheadian distinctions.)

As Simons aptly notes (in Simons 2013), Whitehead's notion of prehension sounds like a form of intentionality. And indeed Whitehead's terminology regularly resonates with mentalistic as well as physicalistic idioms. One thinks of the doctrine of "neutral monism" popular in his day, and specifically of Henri Bergson's metaphysics, to which Whitehead nods approvingly.

Alternatively, we might see prehension as an abstract form that covers three distinct formal relations: the *instantiation* of a form  $P_i$  in  $E^*$ ; the *causation* of  $E^*$  by antecedent events  $E_i$  (temporally prior to  $E^*$ ); and, where  $E^*$  is an event of consciousness, the *intentional apprehension* of  $E_i$  by  $E^*$ . (See Smith 2004, Chapter 7.) In any case, for Whitehead, prehension plays its role in concrecence as considered above.

Arguably, Whitehead's own metaphysical system sought to ground Aristotelian substances in a nexus of concrete events that underlie those substances. For, I believe, Whitehead did not claim to eliminate Aristotelian substances from the world. Rather, he thought concrete "substances" are not basic, but are grounded in a more basic structure of concrete events, or "actual occasions". My point here, in any case, is to note the structure of concrecence and the way in which a concrete entity is *founded*, or *grounded*, in the structure of its concrecence.

## 8.

What I find intriguing in the model of concrecence, as reconstructed here, is how it replaces composition with dependence.

On one traditional view, a concrete object or "substance" is identified with a bundle or *set* of properties, or alternatively a set of property-instances (often called "tropes" nowadays). An improvement on that line is Husserl's treatment of a concrete individual as a *whole* comprising *moments* that are concrete instances of ideal properties. Wholes are not sets: for wholes are defined by the part-whole relation, while sets are defined by the membership relation. Husserl defined moments as *dependent parts* of an individual. But if we separate dependence from parthood, then we find that, for Husserl, a concrete individual is *ontologically dependent* on

a range of property-instances and so ultimately on ideal properties. The individual is not simply a *composite* of property-instances. Rather, the individual is ontologically *dependent* on property-instances: it cannot be that individual—it cannot itself exist—unless those property-instances exist and, per Husserl, are moments of the individual.

Bringing concrescence into Husserl’s story, then, we would find that a concrete individual is *ontologically dependent* on an array of entities (moments) involved in its concrescence. But the individual does not reduce to a set of these entities. Nor is it simply a whole composed of these entities, their mereological sum. Rather, in addition, it derives or *emerges* from them through its concrescence.

We emphasize: The concrescence of a thing is not a structuring of its composition from antecedents, but a structuring of its dependence on antecedents. ... Dependence is not composition!

Now, if we move from concrete substances à la Aristotle or Husserl to concrete events à la Whitehead, we find similarly that a concrete event is *ontologically dependent* on an array of entities involved in its concrescence. Of course an event is not, for Whitehead, identified with the set of events from which it emerges through concrescence. Nor does an event reduce to its role in the “process” of its concrescence. Rather, as an event  $E^*$  comes into being through concrescence, the event  $E^*$  is *ontologically dependent* on its concrescence structure,

$$\langle \langle E_1, \dots, E_n \rangle, \langle P_1, \dots, P_m \rangle \rangle \text{ ---C---} \rightarrow E^*,$$

and so dependent ontologically on the events  $E_i$  and properties  $P_i$  from which  $E^*$  emerges.

According to the theory of concrescence, then, an entity is ontologically dependent on its concrescence structure and so on its antecedents in that structure. And this pattern of dependence defines the basic *mode of being* of the entity—whether the entity is a concrete individual (a “substance”) or a concrete event (an “actual occasion”).

9.

In the concrescence of an entity we see a pattern of *dependence*: the relations of dependence of the entity on its antecedents in its “process” of becoming, hence being, that entity. This structure defines, we said, the basic “mode

of being” of the entity, or a structure encompassing its modes of being. What are “modes of being”?

To clarify the notion, I should like to borrow a page from Roman Ingarden’s *Controversy over the Existence of the World* (Ingarden 2013/1947) (just translated into English). Ingarden was a student of Husserl’s, and his *Controversy* elaborates an intricate theory of modes of being, significantly extending Husserl’s theory of moments, or dependent parts. Of primary relevance here is Ingarden’s focus on how one entity can depend for its being on another.

Ingarden distinguished an entity, its being or existence, and its modes of being. For Husserl, as we noted above, a concrete individual is a whole comprising moments, or dependent parts, that are instances of ideal kinds or properties (such as the particular instance of Red in a red ball). For Ingarden, extending Husserl’s doctrine of moments, an individual is a whole comprising not only qualitative moments (instances of properties or universals), but also “existential moments”, which are instances of its *modes of being*. In particular, where an individual is ontologically dependent on another individual, the former includes a moment that is an instance of that ontological dependence.

As Ingarden argued, a work of art depends for its existence on the consciousness of the artist who created it. By contrast, a tree does not depend for its existence on my consciousness in seeing the tree, or on any other consciousness of the tree. Thus, for Ingarden, an art-work is ontologically dependent on—its being is “relative” to—the artist’s consciousness, while a tree is not ontologically dependent on anyone’s consciousness of the tree. In Ingarden’s ontology, if A is ontologically dependent on B, then (part of) the *mode of being* of A is its way of being dependent on B. And accordingly A includes an existential moment that is an instance of that mode of being-dependent-on-B.

Ingarden’s *Controversy* was aimed at clarifying the issue of realism versus idealism and in consequence correcting what Ingarden took to be Husserl’s radical subjective idealism. I understand Husserl’s “transcendental idealism” as very different from a subjective idealism (see Smith 2013, 161 ff.). In any case, these interpretive issues are beyond our present concerns, as I assume a broadly realist metaphysics. (We shall address Husserl’s account of the “constitution” of objects below.)

With a broadly Ingardenian notion of modes-of-being in mind, let us put dependence in a wider context. We may assume several types or categories of modes-of-being:

- a. Being thus-and-so—following a traditional view of the way universals are instantiated in particulars;
- b. Being actual—setting the scene for traditional modalities of possibility and necessity;
- c. Being possible—assuming metaphysical possibility, as distinct from logical possibility governing inference and representation in a symbolic language;
- d. Being necessary—assuming metaphysical, not logical, necessity;
- e. Being dependent on something—here drawing on the Husserl-Ingarden lines;
- f. Being spatiotemporal—following Husserl’s, and Ingarden’s, conception of being “real”, that is, being in space-time, obviously of relevance to Whitehead’s concerns;
- g. Being ideal (or, as we say today, abstract)—following Husserl’s view and also Ingarden’s of the nontemporal and nonspatial being of both mathematical entities (such as numbers and sets) and meaning entities (sense/Sinn, proposition/Satz, noema, etc.).

We’ll return to the analysis of dependence below. At this point we are focused on its role in the ontology of concrescence: an entity’s mode of being-dependent-on-antecedents in its “process” of concrescence.

Again, on Whiteheadian lines, *becoming* an entity is not a temporal process. The concrescence of an entity traces its ontological dependence on each of various antecedent entities. But this dependence is a *formal*, not a temporal, factor in the formation of the entity. The entity becomes that entity as its dependence upon its antecedents is formed—mathematically formed, as it were.

Concrescence is about the formation of an entity itself. That formation, in a concrescence structure, is not yet a matter of the existence or actuality of the entity. The result of the “process” of becoming is the actual entity itself. And so we go on to posit a variety of modes of being: through its concrescence an entity comes into being, or becomes an *actual* entity. And as it comes into being, its being is dependent on its antecedents. ... That is: following the line of theory we have been sketching.

Given the concrescence of an entity, we move on to its being and modes of being. Thus, if the entity becomes actual, that is a further factor. If it becomes temporal or spatiotemporal, that is a further factor. Thus we cite Whitehead’s principle that “temporality becomes”. If the entity becomes ideal or abstract, that too is another factor. In this spirit we may echo

Ingarden (and more ancient philosophies) in distinguishing the modes of being of an entity from its kinds, properties, relations. These distinctions are distinctions among basic categories of entities in the world—including formal or mathematical structures like concrescence structures, structures of “becoming” an entity.

In the context of such an ontology of modes of being, then, we see dependence as defining the way an entity is dependent for its being on something else. ... The point of this recursion to Ingardenian ontology is to put the ontology of dependence itself front and center stage. In Husserl’s scheme, dependence is primarily set within the doctrine of moments, or dependent parts. Now, with a sketch of an ontology of modes of being, we can sharply separate dependence itself from parthood (in the formation of dependent parts of entities). (Ingarden allows for the separation: Ingarden 2013/1947, 153.)

## 10.

With Husserl’s philosophy in mind, now let us consider how the model of concrescence may apply to the deep structure of consciousness. The Whiteheadian story is a natural fit for consciousness, since consciousness consists in a flow of concrete *events* of consciousness.

Husserl developed detailed analyses of the phenomenological structures of time-consciousness, body-consciousness, and various forms of intentionality—structures that play significantly in everyday experience. Given Husserl’s analyses, we can now see how the deep structure of consciousness involves its ontological dependence on the sense of time, the sense of one’s “lived body” (*Leib*), and the intentional sense of what one sees or thinks or wills. (I assume here my reconstructions of the relevant Husserlian analyses in Smith 2013; again, my concern here is not Husserl exegesis, but adaptation and modification for present purposes.)

On Husserl’s account, consciousness fundamentally involves time-consciousness: not merely consciousness of external things occurring in time, as in seeing the flight of a bird; but “inner” consciousness of time within consciousness. Inner time-consciousness is the form of awareness one has of the temporal flow of one’s own consciousness. The *structure* of that awareness, Husserl proposed, is formed fundamentally in a fusion of “primal” sensory impressions with retentions of just-past impressions and “protentions” of imminent impressions. Some of these sensory impressions

are kinesthetic or proprioceptive sensations, which ground my awareness of my “lived” body in action. Indeed, every act of consciousness occurs within the temporal flow of one’s stream of consciousness, at the base of which is a flow of sensation. So any given act of consciousness—whether wishing or thinking or perceiving or moving volitionally—is *founded* on a basic temporal structure of sensations fused with sensory retentions and protentions. However, an act of perception, such as seeing a pelican gliding over the ocean, is not itself a purely sensory flow. Rather, Husserl holds, the sensory or “hyletic” parts of a perceptual act are informed by interpretive or “noetic” parts of the act. The noetic part(s) of the act entertains an ideal “noematic” sense or meaning, such as “that pelican gliding above the waves”. These sensory and noetic parts of the act are dependent parts, or *moments*, of the act, real-time events constituent of the act, events that—on Husserl’s model of intentionality—draw into the act its ideal, shareable meaning, or noematic content. (See the details in Smith 2013, Chapter 9.)

We can use the model of concrescence now to elaborate the structure of a typical act of consciousness as assayed in Husserl’s phenomenology. Take my seeing that grayish brown pelican gliding just above that cresting wave. If  $V^*$  is my visual experience in this case, we can specify the concrescence structure of  $V^*$  as follows:

$$\langle \langle Si, Ri, Pi \rangle, \langle Ni \rangle, M(Ni) \rangle \text{ —C—} \rightarrow V^*,$$

where:  $V^*$  is my visual experience itself (as characterized above);  $Si$  are its underlying primal sensory impressions,  $Ri$  its underlying sensory retentions,  $Pi$  its underlying sensory protentions;  $Ni$  are its noetic moments, appropriately informing the flow of sensory moments; and  $M(Ni)$  are the noematic meanings entertained by the noetic moments  $Ni$ . The concrescence of  $V^*$  is thus the formal process of  $V^*$ ’s coming into being, that is, becoming  $V^*$ . And by virtue of that process-of-becoming,  $V^*$  is ontologically dependent on that concrescence structure. And so  $V^*$  is dependent on its sensory base (formed from impressions, retentions, and protentions), on its noetic formation (informing the sensory base), and—furthermore—on its ideal noematic meaning-content. (See Smith 2013, Chapter 9, for the fine-grained ontology of noematic contents in relation to acts of consciousness, to which we here add the model of concrescence.)

This concrescence structure underlying an act of consciousness, I submit, is essentially what Husserl characterized as the *constitution* of the act of consciousness, correlated with which is the *constitution* of the noematic

content of the act, and correlatively the *constitution* of the object-of-consciousness as given in the act. Indeed, we may see *genetic* phenomenology as the study of the genesis of consciousness that consists in the concrescence of a particular act of consciousness. As with Whitehead, we then distinguish the temporal reality of consciousness from the ontogenesis, or concrescence, of that temporal reality in the stream of consciousness. A further structure lies in the historical-temporal “genesis” and “sedimentation” of meaning, whereby my seeing that “brown pelican” invokes a concept developed by ornithologists somewhere in the past history of my culture. (Compare Smith 2013, Chapter 9 on the status of noematic meaning.)

11.

Let us return now to the theme of ontological dependence, adapting a variant on Husserl’s notion of “founding”, or *Fundierung*. (Contemporary metaphysicians have recently taken up somewhat related notions of “grounding”: see Correia and Schnieder (eds.) 2012. However, our focus is on *ontological* grounding, not epistemological or explanatory grounding, as we find in some current discussions.)

The simplest form of ontological dependence is defined as follows:

*A depends*, or is *founded*, ontologically on B if and only if  
A could not exist unless B exists  
— i.e., necessarily A exists only if B exists.

Again, A is a *moment* of B if A is a dependent part of B. As we noted above, Husserl focused on property-instances that are, in this sense, moments of an object (e.g., a particular instance of Red in a particular red object). However, I want to focus on the simpler form of dependence, divorced from parthood. For, if we are to say that an event  $E^*$  is ontologically *dependent* on events  $E_i$  in its concrescence structure, those antecedent events  $E_i$  are not *parts* of the actualizing event  $E^*$ . Moreover, we need to allow for different kinds of dependence, defined by different kinds of necessity whereby “necessarily  $E^*$  exists only if  $E_i$  exist”.

In Whitehead’s paradigm, drawn from relativity physics: physical events come into being, or “become”, and thereby take their place in spacetime. In that way a concrete event  $E^*$  is *physically* dependent (ontologically) on its antecedent events  $E_i$ . By contrast, in Husserl’s paradigm, articulated in

his phenomenology but augmented here by the theory of concrescence: an act of consciousness comes into being in the subject's stream of consciousness. And in that way a concrete act of (say) visual consciousness  $V^*$  is *phenomenologically* dependent (ontologically) on its antecedent elements of consciousness (its sensory and noetic bases). By antecedents we here mean ontological antecedents, not temporal antecedents, though Whiteheadian "becoming" brings temporal relations into being through concrescence.

Ontological dependence is dependence *in the being* of something. But what makes an event a physical event is one type of concrescence (according to the laws of physics), and what makes an event an act-of-consciousness is quite a different type of concrescence (according to the principles of phenomenological intentionality theory). And the type of dependence is different in each case.

Accordingly, as we define ontological dependence in the form "necessarily A exists only if B exists", we need to distinguish several different types of possibility and correlative necessity. Philosophers sometimes speak of logical possibility and necessity, governed by rules of logic. Some philosophers speak of conceptual possibility, governed perhaps in a Kantian style by basic conceptual categories such as "object", "property", "number", etc.—where a priori conceptual structures are defined in terms of such basic concepts. Other philosophers speak of metaphysical possibility and necessity, broadly governed by principles such as a theory of universals or essences.

For present purposes, in light of the reflections above, we should distinguish in particular: *physical* possibility, governed by principles of physics; and *phenomenological* or *intentional* possibility, governed by principles of phenomenology, including the theory of intentionality—here following Husserl's model glossed above. Thus, physical dependence is defined in terms of physical necessity, while phenomenological dependence is defined in terms of phenomenological or intentional necessity. How can we explicate the differences between these types of necessity or correlative possibility?

## 12.

Physical possibility or necessity is a matter of what is possible in accordance with the nature of spacetime and matter-energy. Phenomenological

possibility or necessity is a matter of what is possible in accordance with the structure of consciousness, staying with our own forms of conscious experience. Thus, in our current framework we say an act of consciousness is phenomenologically dependent on its base sensory and noetic moments, realized in the subject's stream of consciousness, and thereby further dependent on the noematic content entertained by the act of consciousness. That is to say: the experience cannot exist, with its phenomenological structure, unless that base sensory-noetic structure exists. Since intentionality is central to its phenomenological structure, the experience cannot exist and so be intentionally directed as it is unless its *deep structure* is in place. The deep structure of intentionality is a matter of what we may call intentional possibility—adapting a model devised by Jaakko Hintikka.

In the philosophy of modality, and of modal logic(s), Hintikka has introduced a variety of possible-worlds models for modalities. (See Hintikka 1969.) On Hintikka's model, intentionality amounts to a pattern of directedness of an act of perception or belief (better, judgment), where the act is directed toward an appropriate object or state of affairs in each of a range of "possible worlds". These worlds are not merely logically possible or even metaphysically possible situations. Rather, they are *intentionally possible* situations, small worlds defined as it were in *Lebenswelt* parameters. In Hintikkian terms, an act of visual perception—say, my experience in seeing a pelican gliding over the waves—is directed toward an object or circumstance in each *perceptually possible* world wherein such an object occurs. Similarly, an act of thought or judgment or belief is directed toward a circumstance in each relevant *epistemically possible* world wherein that circumstance occurs.

Consider the visual experience wherein I see that that pelican is gliding over waves. Hintikka's model of the logic of perception is defined for propositional attitudes, assuming perception takes the form of seeing that such-and-such. Then we may say, along Hintikkian lines:

The experience wherein I see that that pelican is gliding over waves is *veridical* if and only if:

for every perceptually possible world compatible with what I see,  
there is before me that pelican gliding over waves.

So the intentionality of the experience consists in a pattern of directedness toward an appropriate situation (that that pelican is gliding over waves) in each *visually possible world* compatible with the content of the experience.

(Here we follow the style of post-Tarskian truth-conditional semantics, transposed here into intentionality theory. Cf. Smith and McIntyre 1982 on this style of analysis of intentionality.)

What type of possibility is perceptual possibility, as opposed to logical or broadly metaphysical possibility? Well, we turn to Husserl's phenomenology for details. On Husserl's model, elaborated with a Hintikkian twist, the "horizon" of my visual experience where I see that pelican gliding over waves consists of the possible circumstances that are "motivated" by the noematic content of my sensory-noetic experience: that is, motivated or "predelineated" given the background meaning on which my experience *depends*. Similarly, a thought or judgment about something is dependent on a horizon of possibilities motivated by the noematic content of that thought. (The assimilation of Husserlian and Hintikkian models is detailed in Smith and McIntyre 1982.)

The phenomenology of perception entails an important ramification of the above veridicality conditions for a sample visual experience. As Husserl elaborated, my visual experience includes my sense of my lived body (*Leib*), in that the object I see before me is presented in relation to my body, as I turn my eyes and head and torso toward the object of vision. Accordingly, the veridicality conditions include the perceived object's spatiotemporal relation to my eyes and my lived body. A *perceptually possible* world for my experience, then, includes my living presence in relation to the intended object. Perceptual possibility is possibility regarding the intended object *in relation to* my lived body.

Different types of intentional or phenomenological possibility and necessity are at work in the concrescence of an act of perception, an act of cogitation, an act of volition in embodied action, and so on. Thus, an act of perception can exist only if its phenomenological bases exist, namely, its sensory and hyletic and noematic "moments". And the act can be directed toward an appropriate object in a variety of perceptually possible situations only if the act is formed by an appropriate form of concrescence. And all this is a matter of phenomenological, not physical, necessity.

13.

The concrescence structure of a particular entity, whether an ephemeral concrete event or an enduring concrete individual, maps out the ways that entity *depends* (ontologically) on its antecedents—not temporal, but

ontological antecedents. We think of these patterns of dependence as *modes of being* characterizing the entity.

Different metaphysical systems focus on different types of such modal characterization. The theory of universals looks to the ways in which a particular *instantiates* properties (kinds, qualities, relations): these are its ways of being thus-and-so, characterized by its instantiating various properties. The theory of causation looks to ways in which a particular event (or object) is *caused* to occur: these are its ways of being brought about, characterized by its being caused in accord with laws of nature. The theory of intentionality looks to the ways an experience is *directed* toward things in the world: these are its ways of being conscious-of-something, characterized by its being directed in virtue of its noematic content or meaning.

Furthermore, the theory of intentionality looks also to the ways an object is “*intended*” in *consciousness* (and so “constituted” phenomenologically). Accordingly, an object, say, the gliding pelican I see, has *modes of being* that define its role in the economy of our subjective experiences, how it is experienced by me or you or our community. These patterns of dependence—running from the object back to our consciousness—do not make the object exist or make it a pelican (as it instantiates biological properties). Rather, these patterns of dependence define its being-experienced—and that “reverse” intentionality is also a part of the deep structure of the object and so of the world. (See Smith 2013, 168–183.)

As Brentano put the point, echoing Medieval thinkers, an object’s being “in-intentio”, its “intentional in-existence”, its being “in consciousness” as it were, is a bona fide feature of the object—a mode of its being. But that view is not an idealist reduction of the object’s being to that of consciousness, say, where the pelican is intended in my visual experience. The Husserlian theory of intentionality allows us to clarify the Brentanian model, now holding that being in the role of object-of-consciousness is a mode of being for the object itself, albeit a mode of being different in structure than its mode of being, say, a pelican. For the pelican plays very different roles in the concrescence structure of the pelican itself (where it is dependent on its antecedent biological properties) and in the concrescence structure of my seeing-the-pelican (where my experience is dependent on its antecedent sensory and noetic moments and the relevant noematic meaning).

Permit me a closing observation, an intuitive and aptly vague hint of a line of re-thinking the deep structures of reality.

The neo-Whiteheadian structure of concrescence may be taken to be a form of *information-processing*, provided we treat “information” in a robust metaphysical form (well beyond configurations of 1’s and 0’s). Antecedent entities  $\langle E_1, \dots, E_n \rangle$  are transformed—formally or mathematically—by an appropriate algorithmic process into a given entity  $E^*$ . We should not see this model as offering a reduction of everything to “information” in our current understanding of digital information, or in the physical implementation of algorithms in silicon chips or what have you. And, again, the “process” defined is not a temporal process running in a computer system, nor even the abstract form of a causal sequence running in real time. The idea is rather that the deep structure of the world involves concrescence, and concrescence can be understood as a formal transformation leading from antecedent entities to the given entity. And that concrescence structure can be understood as carrying “information” forward from antecedent entities to the entity so formed. If you will, “in” each element in the concrescence structure is a “form”, and the total “formation” runs “into” the entities—from  $E_i$  to  $E^*$ —articulated by the formal structure we called a concrescence structure. ... Some physicists today like to say the universe itself is a flow of “information”. In that idea Whitehead might well hear an echo of his theory of ontological “process”, but with Whiteheadian overtones.

Yet this neo-Whiteheading picture of deep structure would not *identify* an entity with its concrescence structure, or with any of its antecedent entities, or even with its role in that structure. Functionalist ontologies want to reduce an entity to its role or function in an appropriate type of process: whether physical, mental, or social roles, as the case may be. Now, the neo-Whiteheadian “informational” theory of concrescence assumes that the entities  $E_i$  and  $E^*$  playing appropriate roles in a concrescence structure have more to them than their roles in the transfer of “information” from  $E_i$  to  $E^*$ . If you like, the ontological “syntax” defining a concrescence structure is enriched by the ontological “semantics” of the entities playing their roles in that structure: the entities  $E_i$  and  $E^*$  carry more “meaning” than their roles in the concrescence. ... And this observation is crucial in the case of an act of consciousness—as long argued in the “semantic” tradition of logical theory (from early Husserl onward).

Importantly, in the case of consciousness, we should not see the concrescence model as reducing consciousness to a flow of digital information. Classical AI still faces the familiar “hard problem” of omitting the lived, phenomenal character of experience, as well as the other hard problem of omitting meaning or noema in favor of syntactic form. Nonetheless, we may envision “information” more broadly, along these neo-Whiteheadian lines, as *forms* in the pattern of concrescence of a thing in the world. Better yet, we may define informational *roles* in the structure of the concrescence of an entity. The entities that play those roles are informational “values” in the algorithmic process that defines the entity at hand. Relevant units of information are the *roles* those entities play in the mathematical process: not the entities that play those roles, but their roles in dependence relations in a pattern of concrescence.

In Husserl’s theory of intentionality, then, we may understand a noematic sense or content as a unit of information in the information-processing model. Accordingly, in the concrescence structure of an act of consciousness (as charted above) we may understand a noematic sense as the type of entity that plays a specific role in the concrescence of an act of consciousness where that noematic sense is entertained in the noetic moment in the act of consciousness. And in that role the noematic sense participates as “information” in the “process(ing)” that defines the act of consciousness. That is: along the lines of the present model of concrescence. (Again, see Smith 2013, Chapter 9, on the fine-grained ontology in Husserl’s theory of noesis and noema.)

Assume the noetic moment in an act of consciousness—carrying the act’s noematic content—is realized in a neural process. Therewith, as the temporal structure of consciousness “becomes”, the temporal flow of “information” in consciousness is founded or dependent on a temporal neurological process. These two aspects of the content of an act of consciousness need to be distinguished: the noetic moment and the neural moment. Similarly, David Chalmers has recently noted the importance of distinguishing the phenomenal and physical aspects of “information”. (See Chalmers 2010, 25ff. on the “double aspect” theory of information.) However, Chalmers’ theory assumes that the information flow at issue is a temporal flow, and at the physical level a causal flow. Our Whiteheadian model, as long stressed, assumes an algorithmic pattern of ontological formation, rather than a temporal and causal flow. I note that Chalmers uses the term “awareness” for a functional property of accessibility to memory, as in a computer model of mind. However, the term “awareness”

naturally refers instead to consciousness's self-awareness: what Brentano called "inner consciousness", what I have accordingly called "inner awareness" (compare Smith 1986). Awareness, or self-consciousness, is then a phenomenal aspect of the "information-processing" we have called the concrescence of an act of consciousness.

15.

In this adventure of ideas I hope we have experienced a creative advance into novelty. The present idea of concrescence comes into being as we (ap)prehend the idea in its own concrescence structure. That structure tracks the idea's dependence on ideas that came into being with Whitehead, Husserl, Ingarden, Hintikka, and other thinkers in the long tradition of metaphysics.

#### REFERENCES

- Aristotle 1995: *Categories*. In: *Aristotle, Selections*, 1995. Translated by Terence Irwin and Gail Fine. Indianapolis: Hackett.
- Chalmers, David J. 2010: *The Character of Consciousness*. Oxford and New York: Oxford University Press.
- Chalmers, David J.; Manley, David; & Wasserman, Ryan (eds.) 2009: *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Clarendon Press.
- Correia, Fabrice, & Schnieder, Benjamin (eds.) 2012: *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge and New York: Cambridge University Press.
- Fiocco, M. Oreste 2014: "Becoming: Temporal, Atemporal and Absolute". Forthcoming in: Nathan L. Oaklander (ed.), *Debates in the Metaphysics of Time*. New York : Continuum/Bloomsbury Press.
- Hintikka, Jaakko 1969: *Models for Modalities: Selected Essays*. Dordrecht and Boston: D. Reidel Publishing Company (now Springer). See especially "On the Logic of Perception", 151–183.
- Husserl, Edmund 2001/1900: *Logical Investigations*, Volumes One and Two. Translated by J. N. Findlay; edited and Revised by Dermot Moran. London and New York: Routledge. (German original, first edition, 1900-01; revised second edition, 1913, 1920.)

- Husserl, Edmund 1983/1913: *Ideas Pertaining to a Pure Phenomenology and a Phenomenological Philosophy, First Book: General Introduction to Pure Phenomenology*. Translated by Fred Kersten. Dordrecht and Boston: Kluwer Academic Publishers, Springer. (German original 1913. Called *Ideas I*.)
- Ingarden, Roman 2013/1947: *Controversy over the Existence of the World*, Volume 1. Translated and annotated by Arthur Szylewicz. Frankfurt am Main and New York: Peter Lang. (Original German edition 1964/65 revised from the Polish editions beginning with 1947/48.)
- Parsons, Charles 2012: *From Kant to Husserl: Selected Essays*. Cambridge, Mass. and London: Harvard University Press.
- Simons, Peter 1987: *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- 1992: *Philosophy and Logic in Central Europe from Bolzano to Tarski: Selected Essays*. Dordrecht and Boston: Kluwer Academic Publishers (now Springer).
- 2013: “Alfred North Whitehead’s *Process and Reality*”. *Topoi*. Published online: 24 April 2013.
- Smith, Barry (ed.) 1992: *Parts and Moments: Studies in Logic and Formal Ontology*. München and Wien: Philosophia Verlag.
- Smith, David Woodruff 1986: “The Structure of (Self-) Consciousness”. *Topoi* 5, 149–156.
- 2002: “Mathematical Form in the World”. *Philosophia Mathematica* (3) Vol. 10, 102–129.
- 2013: *Husserl*. Second edition. London and New York: Routledge. Revised and expanded from the first edition of 2007.
- Smith, David Woodruff, & McIntyre, Ronald 1982: *Husserl and Intentionality*. Dordrecht and Boston: D. Reidel Publishing Company (now Springer).
- Whitehead, Alfred North 1978/1929: *Process and Reality*. Corrected edition, edited by David Ray Griffin & Donald W. Sherburne. New York: The Free Press. (First edition, 1929.)
- Wittgenstein, Ludwig 1974/1921: *Tractatus Logico-Philosophicus*. Translated by David F. Pears & Brian F. McGuinness. London and New York: Routledge. Translated from the German edition of 1921.

PART II  
MIND



ANNEHMEN, PHANTASIEREN UND ENTERTAINING.  
HUSSERL UND MEINONG

Kevin MULLIGAN  
Universität Genf

*Summary*

What are the different forms that imagining may take? Is supposing that  $p$  a form of imagining? Is make-believe fear or desire an exercise of the imagination? What is it to imagine? What is the relation between imagining and “serious” acts or states such as seeing, judging, desiring, fear and belief? Between imagining and what has been called entertaining or grasping a proposition? I reconstruct and evaluate the replies to these questions given by Husserl and Meinong and compare their answers with some later answers.

§1 *Einleitung*

Die von Husserl und Meinong zwischen 1900 und 1902 formulierten Analysen des Annehmens und des Phantasierens gehören zu den einflussreichsten Teilen ihrer frühen philosophischen Arbeiten. Worin bestehen diese Analysen? In welchem Verhältnis stehen sie zueinander? Wie plausibel sind sie? Wie verhalten sie sich zu späteren Versuchen, das Annehmen und das Phantasieren zu verstehen? Im Folgenden versuche ich diese Fragen zumindest teilweise zu beantworten.<sup>1</sup>

---

1. Warum ich mir erlaubt habe, ein Thema aus der Philosophie des Geistes in einer Festschrift für den Ontologen Peter Simons zu behandeln, lässt sich leicht erklären. Im Zentrum von Husserls Ontologie stehen die von ihm genannten *reinen Möglichkeiten*, die ihm zufolge die Gegenstände einer bestimmten Art des Phantasierens sind. Um Husserls Philosophie der Modalität und des Wesens zu verstehen, ist es notwendig, seine Analyse der Phantasie zu verstehen.

§2 *Meinong 1902*

Meinongs bahnbrechende Philosophie der Phantasie, die er 1902 in *Ueber Annahmen* darlegt, kann man mit Hilfe dreier Thesen und einer Vermutung zusammenfassen. Die erste These wird dort angedeutet, wo Meinong sagt, „die Bethätigungen der Phantasie im weitesten Sinne“ machen eine „Hälfte“ der „psychischen Lebensäußerungen“ aus (Meinong 1902, 286),<sup>2</sup> sowie dort, wo er von einer „Zweigetheiltheit“ spricht, „die sich innerhalb der drei Classen [Denken, Fühlen, Begehren] in so verwandter Weise antreffen lässt“ (Meinong 1902, 283). Diese „50:50“ These

MT1      50:50

behauptet also, dass jeder Art von „ernsthaftem“ geistigen Zustand oder Akt<sup>3</sup> sein phantasieähnliches Gegenstück gegenüber steht, und dass alle geistigen Zustände, Episoden oder Akte einer von diesen zwei Arten angehören oder eine Kombination davon sind. Die wichtigsten der von Meinong diskutierten Beispiele sind:

[+ ernsthaft]	[- ernsthaft]
[- Phantasie]	[+ Phantasie]
Urteile	Annahmen Scheinurtheile Phantasieurtheil
Gefühle	Phantasiegefühle Scheingefühle
Begehungen	Phantasiebegehungen Scheinbegehungen <sup>4</sup>

Meinong schreibt:

---

2. Vgl. (Meinong 1977, 384).

3. Meinong spricht in der zweiten Ausgabe von *Über Annahmen* von „Ernsturteilen“ (Meinong 1977, 380), von „Ernstgefühlen“ (316) und von „Ernstbegehungen“ (337).

4. Vgl. (Meinong 1902, 282, 285).

Acceptiren wir vorerst die Benennungen [„Phantasiegefühle“, „Phantasiebegehungen“], so ist nun ohne Weiteres einzusehen, dass diese Phantasiegefühle den wirklichen Gefühlen, die Phantasiebegehungen den wirklichen Begehungen ganz ähnlich gegenüberstehen, wie die Annahmen den Urtheilen : dürfte man jene ganz wohl als Scheingefühle resp. Scheinbegehungen bezeichnen, so nicht minder die Annahmen als Scheinurtheile. (Meinong 1902, 282)

Auch Wünsche sollen ihr Gegenstück haben (Meinong 1902, 238), die man demgemäss „Phantasiewünsche“ nennen könnte, obwohl Meinong den Ausdruck nicht verwendet. Nach Meinong setzen das Urteilen, das Annehmen, das Fühlen, das Begehren und das Wünschen Vorstellungen voraus. Auch hier soll es zumindest eine partielle Zweiteilung geben, z.B. zwischen Wahrnehmungsvorstellungen und Einbildungsvorstellungen:

Zwar beherrscht dieser Gegensatz [zwischen Wahrnehmungsvorstellung und Einbildungsvorstellung],...keineswegs das gesammte Vorstellen: denn von Wahrnehmungs- und daher gegensätzlich hierzu von Einbildungsvorstellungen kann doch nur bei Gegenständen die Rede sein, die ihrer Natur nach ein Wahrgenommenwerden überhaupt gestatten, bei realen Gegenständen also, nicht aber bei idealen Gegenständen wie Aehnlichkeit, Gegensatz u. dgl. (Meinong 1902, 281)

Es ist aber wichtig festzuhalten, dass „Einbildungsvorstellung“ sich zu „Vorstellung“ nicht so verhält wie z. B. „Scheingefühl“ zu „Gefühl“:

Die Erfahrung scheint hier keinerlei Anhaltspunkt dafür zu bieten, dass die Einbildungsvorstellung als Vor- resp. Durchgangsstufe für die Wahrnehmungsvorstellung desselben Gegenstandes anzusprechen wäre, und schon terminologisch ist auffällig, dass der Ausdruck „Vorstellung“ nicht etwa zunächst „Wahrnehmungs Vorstellung“ bedeutet, und man sich daher keineswegs gedrängt fühlt, die Einbildungsvorstellung etwa analog zum „Scheingefühl“ als „Schein Vorstellung“ zu bezeichnen. Vielmehr ist, was eigentlich und zunächst, wenigstens für den ausserwissenschaftlichen Sprachgebrauch, „Vorstellung“ heisst, gerade die Einbildungs Vorstellung, so dass bekanntlich auch viele Theoretiker sich nicht entschliessen können, die Erweiterung der Bedeutung des Wortes Vorstellung zu der des sowohl Einbildungs- als Wahrnehmungsvorstellung einbegreifenden Classennamens mitzumachen. (Meinong 1902, 283)

Was sind nun aber eigentlich Phantasieurtheile, Phantasiegefühle und Phantasiebegehungen? Auf diese Frage gibt Meinongs zweite These eine Antwort:

## MT2 Ein Phantasie-N ist kein N, ähnelt aber einem N

Die Sachlage beim Annehmen, so behauptet Meinong, ist „mit der beim Urtheilen zwar nicht identisch, aber doch in irgend einer Weise verwandt, ihr ähnlich“ (Meinong 1902, 3f.). Es gibt, meint er, ein „Urtheilsartiges ... , das gleichwohl nicht Urtheil ist“. Dies „[legt] die Frage nahe, ob es nicht auch ein Gefühlsartiges geben könnte, das die qualitative Gegensätzlichkeit des Gefühls ebenso aufweist, wie die Annahme die qualitative Gegensätzlichkeit des Urtheils an sich trägt“ (Meinong 1902, 233). Meinongs Antwort auf diese Frage lautet:

Wir haben Thatfachen zu registriren gehabt ..., die ihrer Beschaffenheit wie den Umständen ihres Auftretens nach gefühlsähnlich heißen durften [sic], ohne doch eigentlich das zu sein, was man gewöhnlich sich unter Lust oder Unlust zu denken pflegt: in analoger Weise haben wir den Begehrungen im gewöhnlichen Sinne begehrungsartige Thatfachen an die Seite zu stellen gehabt. Ich habe für Fälle der ersten Art die Bezeichnung „Phantasiegefühle“, für die der zweiten Art den Namen „Phantasiebegehrungen“ vorgeschlagen. (Meinong 1902, 281f.)<sup>5</sup>

Zu den (vermeintlichen) Thatfachen gehört dies:

[D]er urtheilsfähige Romanleser hat für die Personen des Romans ganz ebenso Wünsche, als er Gefühle für sie hat: und diese Wünsche sind streng genommen genau so wenig eigentliche Wünsche, als jene Gefühle eigentliche Gefühle sind. (Meinong 1902, 238)<sup>6</sup>

Was spielt sich im Herzen eines Zuschauers einer Tragödie ab?

[J]ene ‚Furcht‘ und jenes ‚Mitleid‘, oder was sonst die Tragödie zu ‚erwecken‘ die Aufgabe haben mag, was sind sie eigentlich? Eine Furcht, bei der man sich im Grunde doch gar nicht fürchtet, ein Mitleid, das näher besehen eigentlich doch gar nicht ‚weh thut‘, sind das noch ‚Gefühle‘, wie man sie in der Psychologie zunächst zu behandeln pflegt? (Meinong 1902, 234)

[D]er Zuschauer erlebt wirklich etwas in sich. Das ist weder Freude noch Trauer, weder Furcht noch Hoffnung im eigentlichen Sinne, aber etwas Aehnliches, dass jedermann die Anwendung der nämlichen Ausdrücke, wenn

---

5. Einmal schreibt er aber: Ein Phantasiegefühl ist „ein gegenwärtiges, in seiner Weise ganz wirkliches Gefühl“ (Meinong 1902, 239).

6. Goethe wäre nicht einverstanden – vgl. (Eckermann 1836-1848, Aufz. v. 21. 7. 1827).

er sie auch sofort als uneigentlich verspürt, doch ohne Weiteres versteht. (Meinong 1902, 236f.)<sup>7</sup>

Auch Wittgenstein unterscheidet in einem ähnlichen Zusammenhang zwischen zwei Anwendungen eines Ausdrucks:

Ja; wir sagen von Leblosem, es habe Schmerzen: im Spiel mit Puppen z.B. Aber diese Verwendung des Schmerzbegriffs ist eine sekundäre. (Wittgenstein 1977, §282)

Man sagt nur von solchen Kindern, sie spielen Eisenbahn, die von einer wirklichen Eisenbahn wissen. Und das Wort Eisenbahn im Ausdruck ‚Eisenbahn spielen‘ ist nicht bildlich gebraucht, oder im übertragenen Sinn. (Wittgenstein 1984, 452f., §800)<sup>8</sup>

Meinong bezieht sich ebenfalls auf das Spiel von Kindern und „spricht von der „Sicherheit“, mit der diese „bereits in frühen Jahren Spiel und Ernst zu unterscheiden wissen“. Das spielende Kind, so Meinong, befindet sich nicht im „Zustand der Täuschung“ (Meinong 1902, 41).

Meinong verwendet eine Reihe von Umschreibungen, um das Phantasieren in den Griff zu bekommen. Eine Unterart der Phantasiegefühle, nämlich affektive Zustände, die die Annahme voraussetzen, dass etwas existiere oder nicht existiere, nennt er „Quasigefühle“:

[Es ist eine] Thatsache, dass die beiden Objective Existenz und Nichtexistenz nicht nur durch das Urtheil sondern auch durch die Annahme erfasst werden können, und demgemäss die Werthgefühle sich nicht nur als Urtheilsgefühle sondern eventuell auch als Annahmefühle oder genauer gefühlsartige Zustände darstellen, durch die das Subject dort auf Annahmen reagirt, wo im Urtheilsfalle sich die Werthhaltungen einzustellen pflegen. Diesen Quasigefühlen gegenüber, in denen wir einen speciellen Fall dessen haben erkennen müssen, was oben als „Phantasiegefühl“ zu benennen sich empfahl, ... (Meinong 1902, 250f.)

1910 spricht Meinong auch von „etwas Gefühlartige[m]..., das noch kein volles Gefühl ist“ (Meinong 1977, 312). Generell verhält sich, wer phantasiert, „als ob“ etwas der Fall wäre:

---

7. Wenn alle Gefühle, die der Leser von *Hamlet* „bezüglich“ Ophelia hat, uneigentliche Gefühle in diesem Sinne sein sollten, dann verschwindet das „paradox of fiction“.

8. Zu den Unterscheidungen zwischen „primärer“ und „sekundärer“ Bedeutung bei Meinong und Wittgenstein vgl. (Mulligan 2012, 115–124).

Die Zumuthung, auch nur vorübergehend ein Märchen zu glauben, wird sich wenigstens der „Gebildete“ nur sehr ungern bieten lassen. Dagegen wird er vielleicht ohne sonderliches Widerstreben einräumen, dass, obwohl er den Roman, den er eben liest, doch in der Regel auch nicht für eine „wahre Geschichte“ nimmt, er sich während der Lektüre desselben zu der Handlung und zu den einzelnen Personen recht ähnlich verhält, als ob sie wirklich wären. (Meinong 1902, 57)

Oder man kann auch sagen, er „fingiere“:

[Es] wird ... nun auch Niemandem schwer fallen, sich daran zu erinnern, wie oft er ... Aeusserungen [gemacht hat] wie die, dass der Spielende an sich und Anderen Eigenschaften, Situationen u. dgl. „fingiere“, um dann häufig, so lange das Spiel währt, zu thun, als ob er an die Fiction glaubte, obwohl ihm solches völlig ferne liegt. (Meinong 1902, 42)

Sogar das Schliessen soll mit fingierten Urteilen oder Annahmen auskommen können:

Ja man legt bei Feststellung der beim Schliessen in Frage kommenden Gesetzmässigkeiten augenscheinlich gar keinen Werth darauf, mit wirklichen Urtheilen als Prämissen und Conclusionen zu operiren, indem man alles Erforderliche ganz wohl an fingirten Urtheilen einzusehen vermag, deren Richtigkeit man ohne Schaden in suspenso lässt, falls man nicht etwa gar von ihrer „materiellen“ Falschheit überzeugt ist. (Meinong 1902, 76f.)<sup>9</sup>

Die Ausdrücke, die Meinong einführt beim Versuch, das Wesen des Phantasierens zu umschreiben – „quasi“, „als ob“, „fingieren“ usw. –, wurden im Laufe des zwanzigsten Jahrhunderts immer wieder für das Philosophieren über das Phantasieren herangezogen. Insbesondere Autoren, die das Phantasieren mittels Ausdrücken wie „make believe“, „pretence“ oder „simulation“ zu verstehen versuchen, machen von ihnen Gebrauch. Helfen solche Ausdrücke uns aber wirklich weiter? Einer ihrer Nachteile ist, dass sie gern verwendet werden, um unser Verhalten und unsere Handlungen zu beschreiben, und dabei oft die Konnotation der absichtlichen Täuschung haben – „er tat nur so, als ob er sie hasste“.

---

9. Meinongs Untersuchungen der Rolle der Annahmen im Schliessen (vgl. dazu Heller 1929), die er 1910 zum Teil revidiert, nehmen einige Thesen der Theoretiker der „natürlichen Deduktion“, Jaśkowski und Gentzen, sowie von Ryle, vorweg. In seiner Diskussion einer Methode einer „arbitrary supposition“ bezieht sich Jaśkowski auf Łukasiewicz, der bei Meinong gehört hat (Jaśkowski 1967, 232); Twardowski und andere polnische Denker haben Meinongs Kategorie der Annahmen gründlich diskutiert. Es kann sein, dass Gentzen Meinongs Annahmehlehre aus den Schriften von Hertz kannte (vgl. Hertz 1931, 377; 1935, 235).

Meinongs Ausführungen über die bloße Ähnlichkeit zwischen ernsthaften Akten und Phantasieakten erinnern an ein Phänomen, das oft von Brentanos Erben diskutiert wurde. Das Verhältnis zwischen einem Phantasieakt und seinem ernsthaften Gegenstück scheint dasselbe zu sein, wie das Verhältnis zwischen falschem Geld und Geld. Falsches Geld ist kein Geld, hat aber einige Eigenschaften, wie beispielweise das Aussehen, mit Geld gemeinsam. Das Wort „falsch“ im Ausdruck „falsches Geld“ determiniert nicht, sondern modifiziert das Nomen „Geld“.

Meinongs Versuch, die bloße Ähnlichkeit zwischen ernsthaften Akten und ihren phantasieartigen Gegenständen zu verstehen, hat zwei Aspekte. Zum einen werden Phantasieakte teilweise negativ bestimmt: wer annimmt, dass *p*, urteilt *ohne* zu glauben, dass *p*. Darauf werden wir gleich zurückkommen. Zum anderen gibt es für Phantasieakte auch die folgende positive Bestimmung:

MT3      Das Phantasieren ist direkt dem Wollen unterworfen

Soweit ich sehen kann, formuliert Meinong diese These allerdings nicht ausdrücklich. Sie wird aber zumindest nahegelegt durch die Behauptung, dass, wer phantasiere, auch fingiere. Aus Meinongs Prinzip „unbeschränkter Annahmefreiheit“ scheint zumindest „in betreff isolierter Annahmen“ zu folgen, dass das Annehmen direkt dem Wollen unterworfen ist. (Meinong 1902, 261). Diese Annahmefreiheit ist eine Voraussetzung derjenigen epistemischen Autorisierung, die die Grundlage der Logik des natürlichen Schliessens bildet:

The first rule of derivation to be introduced is the *rule of assumptions*....This rule permits us to introduce at *any* stage of an argument *any* proposition we choose as an assumption of the argument. (Lemmon 1965, 9)

Eine Version der These (MT3) findet man bei Meinongs genialem Schüler, Stephan Witasek:

Wir können unsere Überzeugung nicht nach unserer Willkür modeln; im Aufstellen von Annahmen dagegen hat unser Wollen durchaus freie Hand, da sind wir nicht beschränkt durch Tatsachen und Meinungen...Deshalb stellt das Gebiet der Annahme auch vorzugsweise ein Gebiet der Phantasiebetätigung dar ... (Witasek 1923, 311)

Auch „die Phantasiegefühle“ und die „Phantasievorstellungen“ sind laut Witasek „dem Willen bis zu gewissem Grade dienstbar“ (Witasek 1904,

135).<sup>10</sup> Es stimmt zwar, dass Termini wie „als ob“ und „fingieren“ meistens im Zusammenhang mit einer körperlichen Handlung oder einem körperlichen Verhalten verwendet werden. Aber es gibt auch geistige oder psychologische Tätigkeiten, auf die sie zutreffen – nämlich die verschiedenen Arten des Phantasierens. So wie ich mich entscheiden kann, gewisse Körperteile in Bewegung zu setzen, kann ich mich auch entscheiden zu phantasieren. Im Gegensatz zum Sehen, zum Urteilen und zum Wollen selbst (es gibt kein Wollen-Wollen) ist das Phantasieren wie das Bewegen von Körperteilen direkt dem Wollen unterworfen. Was es heisst, *direkt* dem Wollen unterworfen zu sein, lässt sich besser verstehen, wenn man sich daran erinnert, dass z. B. das Sehen und das Geniessen relativ leicht zu kontrollieren sind, obwohl sie nur *indirekt* dem Wollen unterworfen sind. Im ersten Fall reicht es meistens, die Augen aufzumachen, im zweiten eine Flasche Château Margaux zu trinken. Sind nun Phantasiegefühle direkt dem Wollen unterworfen, wie etwa das visuelle Phantasieren oder das Annehmen, dass es keine EU gäbe? Witasek scheint dies zu verneinen. Und wer meint, es brauche gewisse Requisiten, wie z.B. Kunstwerke, um Phantasiegefühle zu erzeugen, wird Witasek vermutlich recht geben. Denn dann wären Phantasiegefühle, ähnlich wie das Sehen, relativ leicht zu kontrollieren, aber dennoch nicht direkt dem Wollen unterworfen.

Betrachten wir nun die negative Bestimmung in Meinongs Versuch, die Ähnlichkeit zwischen ernsthaften Akten und ihren Gegenständen zu verstehen. Die „Urteilsähnlichkeit bei der Annahme“ formuliert er so: „Annahme ist Urtheil ohne Ueberzeugung“ (Meinong 1902, 257).<sup>11</sup> Da das Urteilen etwas ist, was vorkommt, also episodisch ist, und da die Überzeugung etwas ist, was dauert (egal wie lang die Dauer ist), also zuständig ist, sollte man Meinongs Behauptung wohl so verstehen: wer zum Zeitpunkt  $t$  annimmt, dass  $p$ , urteilt zu  $t$ , dass  $p$  ohne gleichzeitig zu glauben, dass  $p$ . Das aber ist unmöglich.<sup>12</sup> Meinong *scheint* dies auch zu wissen, wenn er sagt, eine Annahme sei „urteilsähnlich“. Andererseits sagt er aber auch, man könne gut behaupten: „Urtheil ist Annahme unter Hinzutritt der Überzeugung“ (Meinong 1902, 257). Wenn das Annehmen, dass  $p$ , die Überzeugung, dass  $p$  ausschliessen soll, dann muss man das Prinzip der Annahmefreiheit aufgeben.<sup>13</sup> Es stimmt aber nicht, dass das Annehmen die Überzeugung ausschliesst; wer glaubt, das Belgische

10. Vgl. a. a. O., 215f. Nach Witasek (1904, 24) ist das Annehmen auch ein Fingieren.

11. Vgl. a. a. O., 2f., 277.

12. Vgl. (Reinach 1989, Bd. 1, 99f.).

13. Vgl. (Künne 1995, 371f.).

Kaiserreich existiere, kann sehr wohl annehmen, dies sei nicht der Fall, um Schlüsse aus dieser Annahme zu ziehen.<sup>14</sup>

Unter Beibehaltung seiner irreführenden Formulierung fragt sich Meinong, ob das Phantasiegefühl sich zum Gefühl wie die Annahme zu einem Urteil ohne Überzeugung verhält. 1910 schreibt unser gewissenhafter österreichischer Beamter:

Das Analogon [für das Phantasiegefühl] zu dem, was als "Überzeugungsmoment" durch sein Hinzutreten zur Annahme aus dieser gleichsam das Urteil macht, weiss ich freilich zurzeit nicht anzugeben und kann nur hoffen, dass künftige Analyse hierin erfolgreicher sein wird. (Meinong 1977, 313)

In welchem Verhältnis stehen Meinongs positive und negative Bestimmungen des Phantasierens zueinander? Seine negative Bestimmung der Annahme erlaubt ihm, die Ähnlichkeit zwischen Urteilen und Annehmen im Sinne der sogenannten „Ähnlichkeit durch gleiche Teile“ zu verstehen (Meinong 1902, 279). Das Annehmen enthält abstrakte Teile oder Momente, die man auch beim Urteilen findet, nämlich Affirmation oder Negation und Inhalt. Auch beziehen sich sowohl das Urteilen als auch das Annehmen auf Objektive (Sachverhalte). Dem Annehmen *fehlt* aber das Moment der Überzeugung. Wie wir gerade gesehen haben, konnte Meinong diesen Ansatz jedoch nicht verallgemeinern. Die These, wonach das Phantasieren direkt dem Wollen unterworfen sei, schaut da vielversprechender aus. Damit gäbe es immer noch eine Reihe von Ähnlichkeiten zwischen dem Sehen und dem visuellen Phantasieren, zwischen Annehmen und Urteilen usw., aber die in Frage kommenden Ähnlichkeiten wären dann keine „Ähnlichkeiten durch gleiche Teile“ sondern einfache, nackte oder „brute“ Ähnlichkeiten.<sup>15</sup>

---

14. Die Überzeugung oder der Glaube steht, wie gesagt, zum Urteilen wie ein Zustand zu einer Episode. Das Phantasieren hingegen scheint kein zuständliches Gegenstück zu haben, obwohl es Phantasiedispositionen wie auch Urteilsdispositionen gibt (Mulligan 1989, 136). Künne weist darauf hin, dass es im System unserer alltagspsychologischen Kategorien eine „merkwürdige terminologische Lücke“ gibt: „It seems to me that we have no well-established succinct title for that state which is related to believing that *p* as merely entertaining the thought that *p* is related to judging that *p*. You are in such a state with respect to the components of a disjunction, for example, if you believe the disjunction without believing either of its disjuncts. It is the state you are in with respect to (the proposition that) *p* if and only if you *don't believe* that *p* but are *capable of entertaining* the thought that *p*“ (Künne 1996, 62). Man behauptet oft, Glauben, Gefühle und Begehungen weisen Grade auf (vgl. Mulligan 2013). Das Phantasieren aber scheint nicht steigerungsfähig zu sein (im Gegensatz etwa zur *Deutlichkeit* des Phantasierens), wohl deshalb, weil das Phantasieren eine Tätigkeit ist (vgl. Gerrans & Mulligan 2013).

15. Vgl. (Mulligan 1999).

Um Meinongs Behauptung besser zu verstehen, dass ein Urteil eine Annahme unter Hinzutritt von Überzeugung sei, empfiehlt es sich, eine seiner Vermutungen zu zitieren, die auch für den Vergleich mit Husserl wichtig sein wird:

MV1 „[D]ie Urtheile [geben] in irgend einem Sinne eine Art Oberstufe zu den Annahmen als Unterstufe [ab]“. (Meinong 1902, 280)

Bezüglich dieser Vermutung stellt Meinong sich 1902 zwei Fragen, die er „derzeit ausser Stande“ sei, „bündig zu beantworten“:

[1] [H]aben wir nicht im Urtheile insofern einen complexen Thatbestand vor uns, dass darin jedesmal eine Annahme enthalten ist und dann nur noch etwas dazu, das wir als Ueberzeugungsmoment kennen? ...

[2] ob wir in der Annahme nicht etwa einfach eine Vorstufe des Urtheils vor uns haben, die man beim Concipiren des Urtheils normalerweise jedesmal passiren müsste, und die dafür, wenn erreicht, auch allemal eine gewisse Urtheilchance repräsentirte ... (Meinong 1902, 279)

„Im Ganzen steht es“, meint er aber dann (und wiederholt dies 1910), „für eine affirmative Beantwortung beider obigen Fragen keineswegs ungünstig“ (Meinong 1902, 280). Seine Vermutung lässt sich wie folgt ausdrücken:

Wenn eine Annahme ein Urteil ohne Überzeugung ist, dann ist eine Annahme ein Urteil ohne Überzeugung, *weil* ein Urteil eine Annahme unter Hinzutritt der Überzeugung ist.

Mit anderen Worten, von den beiden, Annehmen und Urteilen, ist es das Annehmen, das die Hosen an hat.

Um Meinongs Beschreibung des Phantasierens besser zu verstehen, empfiehlt es sich, die bisherigen Ausführungen mit Meinongs analysierender Beschreibung der Intentionalität im Allgemeinen zu konfrontieren. Wie andere Erben Brentanos unterscheidet Meinong dabei zwischen Aktqualitäten oder Modi, Inhalten und Gegenständen oder Objektiven. Er übernimmt auch die These, dass Gefühle, Begehren, Wollen und Wünschen eine positive oder negative Valenz haben (Meinong 1902, 282). Und obwohl Meinong Brentanos merkwürdige Auffassung, wonach die Intentionalität des Urteilens nicht propositional sei, aufgibt, hält er an Brentanos ebenso merkwürdiger Ansicht fest, das Urteilen selber – nicht der Inhalt des Urteilens – sei positiv oder negativ.<sup>16</sup> Er meint, die Annahme habe eine

16. Vgl. dazu (Mulligan 1989).

„bejahende“ oder „verneinende“ Qualität, ganz so wie auch die Qualität des Urteilens immer positiv oder negativ sei (Meinong 1902, 119).<sup>17</sup>

Meinongs falsche These, eine Annahme sei ein Urteil ohne Überzeugung, hängt mit seiner unplausiblen „Oberstufen“-Theorie zusammen, wonach ein Urteil eine Annahme unter Hinzutritt von Überzeugung sei. Wenn Meinong damit recht hätte, dann könnte ein propositionaler Inhalt in einem Subjekt gleichzeitig durch die zwei verschiedenen Aktqualitäten des Annehmens und des Urteilens qualifiziert sein. Wie wir sehen werden, wird genau diese Möglichkeit von Husserl abgelehnt. Die Aktqualität der Überzeugung wie auch die Aktqualität des Urteilens verlangen einen propositionalen Inhalt, sind aber nicht auf einer Aktqualität des Annehmens „fundiert“.

Meinong scheint das Phantasieren als eine Aktqualität zu verstehen, die im Falle des Annehmens und des Phantasiegefühls eine positive oder negative Valenz hat. Hinzu kommt, dass es für ihn kein Phantasieren *tout court* gibt. Wer phantasiert ist immer Träger eines Phantasiegefühls, eines Annehmens, eines Phantasiebegehrens oder eines Phantasiewünschens (wenn überhaupt etwas, dann könnten allein die Einbildungsvorstellungen Beispiele eines Phantasierens *tout court* sein). Diese Deutung lässt sich auf Englisch folgendermassen ausdrücken:

- T1 If  $x$  imagines, then  $x$  imagines  $y$  or  $x$  imagines that  $p$  or  $x$  imagines being  $F$
- T2 If  $x$  imagines  $y$  or  $x$  imagines that  $p$ , then  $x$  imagines  $G$ -ing  
If  $x$  imagines  $y$ , then  $x$  imagines seeing/hearing/touching/  
smelling/emoting (admiring, being pleased by ...) ...  $y$   
If  $x$  imagines that  $p$ , then  $x$  imagines judging/desiring/regretting/  
being pleased/afraid ... that  $p$

Wie soll man „imagines seeing“ oder „imagines judging“ verstehen? Man könnte der Meinung sein, „imagines“ funktioniere hier wie „enjoyed“ in „ $x$  enjoyed listening to the music“ oder „ $x$  enjoyed meeting  $y$ “. Wenn Sam es genoss, der Musik zuzuhören („enjoyed listening to the music“), dann gab es ein Zuhören, nämlich das Zuhören, das er genoss. Wenn Sam jedoch „imagined seeing  $y$ “, dann gab es kein Sehen, das er phantasierte oder sich vorstellte. Genau gesagt müsste „to imagine seeing“ als „to imagine-seeing“ wiedergegeben werden, „to imagine judging“ als „to imagine-judging“.

---

17. Vgl. a.a.O., 280, 258.

„to imagine desiring“ als „to imagine-desiring“ usw. Mit anderen Worten: das Phantasieren gehört gänzlich auf die Seite der Aktqualität. Diesem Umstand lässt sich im Deutschen dadurch Rechnung tragen, dass man „imagines-seeing“ als „sieht<sub>ph</sub>“ und „imagines-desiring“ als „begehrt<sub>ph</sub>“ usw. transkribiert. Damit wird deutlich, dass das *Sehen<sub>ph</sub>* zum *Sehen*, das *Begehren<sub>ph</sub>* zum *Begehren* usw. in der Beziehung der einfachen, nackten oder „brute“ Ähnlichkeit steht. Das heisst, diese Aktqualitäten oder Modi haben keine gemeinsamen Teile oder Momente.<sup>18</sup>

### §3 *Martin 2002*

Gewisse Bestandteile von Meinongs Analyse des Phantasierens und Annehmens lassen sich besser verstehen, wenn wir sie mit Martins Theorie der sinnlichen Phantasie vergleichen:

By sensory imagining I have in mind those distinctive episodes of imagining or imaging which correspond to our use of the distinct senses: so we talk of visualising corresponding to seeing, or listening in one's head parallel to audition, and so on. (Martin 2002, 403)

---

18. Meinongs *Philosophie der Phantasiegefühle und der Phantasiebegehrenen* wurde von seinen Schülern – Saxinger, Schwarz, Benussi und Witasek – weiter entwickelt. Vgl dazu (Smith 1985) und (Vendrell Ferran 2009, 2010). Witasek formuliert einen wichtigen Einwand gegen Meinongs Theorie der Phantasiegefühle, auf den Meinong mehrere Male zurückkommt. Meinong zufolge ist ein Phantasiegefühl gänzlich phantasieartig: Wer „Furcht“ erlebt aufgrund einer visuellen Phantasie eines Hundes, fürchtet sich<sub>ph</sub> und sieht<sub>ph</sub>. Warum, so fragt Witasek (1904, 118), soll diese „Furcht“ aber nicht wirkliche, ernsthafte Furcht auf der Grundlage einer visuellen Phantasie sein? Das Vorhandensein einer visuellen Phantasie reicht doch aus, um zu erklären, warum der, der „sich fürchtet“, nicht flieht. Witaseks Herausforderung ist für Meinong deshalb besonders schwierig zu beantworten, da dieser Lust und Unlust nicht als affektive Aktqualitäten betrachtet sondern als Bestandteil des *Inhalts* jeden Gefühls. So aufgefasst gleichen Meinongs Lust und Unlust mehr den sogenannten Gefühlsempfindungen als den Gefühlen. Freunde der Phantasiegefühle und der „make-believe emotions“ geben meistens zu, und zwar aus guten Gründen, dass die zu einem Phantasiegefühl gehörigen körperlichen Empfindungen, wie etwa ein Prickeln auf der Haut, nicht phantasieartig, sondern ganz ernsthaft sind. Wenn man Meinongs These, ein Phantasiegefühl sei gänzlich phantasieartig, auf die in Frage kommenden Aktqualitäten beschränkt, dann könnte man sagen, die Lust und Unlust, die den Inhalt eines Phantasiegefühls ausmachen, seien wirkliche Lust und Unlust. – Eine interessante *Anwendung* der Kategorie der Phantasiegefühle liefert Meinongs Analyse der direkten Erinnerung an Gefühle: man erfasst ein vergangenes Ernstgefühl mittels eines Phantasiegefühls (Meinong 1973, §16, 443f.). Zu den Phantasiegefühlen und ihren Rollen vgl. auch (Mulligan 2009, Gerrans & Mulligan 2013).

Für unsere Zwecke relevant ist folgende These, von Martin „Abhängigkeits-These“ genannt:

The Dependency Thesis, as I shall call it, claims ...: to imagine sensorily a  $\varphi$  is to imagine experiencing a  $\varphi$ . On this view, one kind of phenomenally conscious state, an event of imagining, takes as its object another type of conscious state of mind, a sensory experience. (Martin 2002, 404)

Der zweite Satz behauptet, dass das „imagining experiencing“ aufzufassen sei als ein Phantasieren (Sich-Vorstellen), dessen *Gegenstand* ein Erleben sei. Der Gegenstand des Phantasierens beschränkt sich jedoch nicht allein auf ein Erleben, sondern schliesst auch den Gegenstand dieses Erlebens, eine Szene oder Situation, ein:

Likewise I assume at this stage that Dependency commits one to the claim that in imagining some scene one thereby imagines an experience of the scene – it is no part of Dependency to deny that one imagines the scene when one imagines an experience of the scene. (Martin 2002, 404)

Martins These bezieht sich auf alle Arten der sinnlichen Phantasie:

I suggest that the Dependency Thesis is plausibly taken as applying to all sensory imaginings and all sensory experiences, at least for those cases where one imagines a situation ‘from the inside’. (Martin 2002, 404)

Martin erwähnt dann eine alternative Auffassung, „the community view“, die einem Teil der Analyse, die wir Meinong zugeschrieben haben, entspricht:

One approach here would be to suppose that the similarities between the two kinds of state [sensory experiences, sensory imaginings] are to be explained in terms of community of properties between them. Sensory experiences possess certain phenomenal properties and are thereby sensuous experiences, so sensory imaginings echo this aspect of sensuous experience by having some of the same properties, although possibly to a different degree. One will get different versions of the community view, depending on one’s conception of the nature of phenomenal properties. (Martin 2002, 405)

Er vergleicht nun die „community view“ mit seiner eigenen Ansicht:

But the community of properties is not the only way in which the intuitive similarity of imagining and sense experiencing could be accounted for. Instead, one may think of the connection as an internal or *intentional* one: sensory imagining is experiential or phenomenal precisely because *what is*

*imagined* is experiential or phenomenal ... (Martin 2002, 406; Hervorhebung K.M.)

Sowohl für Meinong als auch für Martin ist also „to imagine sensorily to imagine experiencing“. Beide kennen die Ansicht, wonach das sinnliche Phantasieren und das sinnliche Erleben einige Eigenschaften gemeinsam haben. Martin zufolge ist „visualising“ „the imagining of visual experience. When I visualise an apple, I imagine it through imagining a visual experience as of an apple“ (Martin 2002, 410f.; Hervorhebung K.M.). Daraus folgt, dass alle Unterschiede zwischen den Arten des sinnlichen Phantasierens – das visuelle, auditive, taktile, kinästhetische, propriozeptive ... Phantasieren – beim Gegenstand des Phantasierens liegen. Dagegen lässt das sinnliche Phantasieren à la Meinong zwei Spezifikationen zu: auf der Seite des Gegenstandes *und* auf der Seite der Aktqualität: wer sinnlich phantasiert, sieht<sub>ph</sub>, hört<sub>ph</sub>, oder berührt<sub>ph</sub> usw.<sup>19</sup>

### §3 Husserl 1900–1901

Kommen wir jetzt zu Husserls Phantasien. Wer Husserls Analyse der Phantasie kennenlernen will, muss sich zuerst durch einen terminologischen Dschungel kämpfen und Begriffe wie

nicht-setzende Akte, Aktqualität, Aktmaterie, Gegenwärtigung, Vergewärtigung, Neutralitäts-Modifikation, qualitative konforme Modifikation, imaginative konforme Modifikation, Phantasievorstellungen, Bildbewusstsein ...

verstehen.<sup>20</sup> Bevor wir zu den Sachen selbst kommen, müssen wir also zumindest Husserls erste Taxonomie des Geistes kennenlernen.

Gemäß Husserls erster ausführlicher Analyse der Struktur von sogenannten „Akten“ oder „intentionalen Erlebnissen“ (Husserl 1984, V §30, 453) sind diese entweder *objektivierend* – Sam urteilt, dass es regnet; Sam sieht Maria – oder *nicht-objektivierend* – Sam bewundert Maria, freut sich darüber, dass Maria angekommen ist oder wünscht, dass Maria nie

---

19. Ich übergehe hier eine Disanalogie zwischen den Auffassungen von Martin und Meinong. Martin behandelt das nicht-propositionale sinnliche Phantasieren. Meinong fasst die Wahrnehmung und infolgedessen auch die Phantasiewahrnehmung propositional auf.

20. Vgl. auch (Mulligan 1999).

ankommen möge. Worin besteht der Unterschied? Husserl sagt dazu nur, dass nicht-objektivierende Akte objektivierende Akte voraussetzen, einfache objektivierende Akte hingegen nicht-objektivierende Akte nicht voraussetzen könnten (Husserl 1984a, VI §27, 624; 1984, V §41, 493). Zu jedem Akt gehört eine Aktqualität und eine Aktmaterie (Husserl 1984, V §22, 426). Die Aktmaterie eines nicht-objektivierenden Aktes ist die Materie des zugrundeliegenden objektivierenden Aktes: die Materie von Sams Freude, [Maria ist angekommen], ist auch die Materie seiner Überzeugung (Husserl 1984, V §41, 494). Wie die Akte selbst sind die dazugehörigen Materien entweder propositional oder nicht propositional. Nicht-propositionalen Materien sind entweder begrifflich (sie entsprechen einfachen oder komplexen Bedeutungen oder Begriffen) oder nicht-begrifflich, wie im Falle der einfachsten Art visueller Wahrnehmung (Husserl hat anfänglich alle nicht-propositionalen Materien „nominal“ genannt, eine irreführende Terminologie, die er später fallen liess). „Aktqualitäten“ werden die Modi genannt, die Materien oder Inhalte qualifizieren und den unterschiedlichen psychologischen Verben oder Adjektiven entsprechen, zum Beispiel: etwas *sehen*; *sehen*, dass *p*; etwas *bewundern*; *traurig sein*, dass *p*; *urteilen*, dass *p*; *annehmen*, dass *p*; *wollen*, dass *p*; *wünschen*, dass *p*; *sich freuen*, dass *p*<sup>21</sup>. Einige Aktqualitäten sind Stellungnahmen: „Stellungnahmen‘ des Vermutens, des Wünschens usw.“ (Husserl 1984, V §32, 472).

Nicht-objektivierende Akte sind entweder volitiv oder affektiv. Beispiele der zweiten Kategorie sind episodische Gefühle wie das „Gefallen an einer Melodie“, das „Missfallen an einem schrillen Pfiff“, oder die „Freude oder Unfreude über etwas“ (Husserl 1984, V §15(a), 402). Die Intentionalität dieser Gefühle ist nicht autonom, sondern in der Intentionalität der ihnen unterliegenden objektivierenden Akte gegründet:

Sie... ‚verdanken‘ ihre intentionale Beziehung gewissen ihnen unterliegenden Vorstellungen. Aber im Sinn der Rede vom Verdanken liegt ja ganz richtig, dass sie selbst nun auch das *haben*, was sie den anderen verdanken. (Husserl 1984, V §15(a), 404)

Was genau sind nun aber Aktqualitäten? Der obigen Antwort zufolge sind z. B. das Wollen, das Urteilen und das Annehmen Aktqualitäten. Husserl

---

21. Die Wünschqualität (Husserl 1984, V §30, 453, V §32, 474); die Qualität der Freude (Husserl 1984, V §42, 495); die Qualität des Gefallens (Husserl 1984, V §42, 497); die Qualität der Fürwahrhaltung (Husserl 1984, V §42, 495); die Qualität des Urteilens (Husserl 1984, V §32, 474) und die Qualität des Fragens (Husserl 1984, V §32, 474).

scheint allerdings auch eine zweite Antwort zu geben, wenn er vom Gebiet der objektivierenden Akte spricht:

Wir unterschieden innerhalb der nominalen Akten die setzenden und die nicht-setzenden ... Die Urteile als setzende propositionale Akte haben also ihre Korrelate in blossen Vorstellungen als nichtsetzenden propositionalen Akten. (Husserl 1984, V §38, 499)

Wie verhalten sich diese zwei Charakterisierungen von Aktqualitäten zueinander? Die erste erwähnt Aktqualitäten von objektivierenden und von nicht-objektivierenden Akten, die zweite erwähnt nur die Qualitäten von objektivierenden Akten. Die erste kommt mit nominalisierten Verben für „Erlebnisse“ aus. Die zweite verwendet einen eher technischen Ausdruck, nämlich „setzend“ und eine nur negative Charakterisierung („nicht-setzend“). Kann eine positive Qualität rein negativ charakterisiert werden? Husserl kombiniert die zwei Charakterisierungen in der folgenden Liste:

Die Qualität der Akte, die Weisen des Glaubens, blossen Dahingestelltseinlassens, Wünschens, Zweifelns usw. ... (Husserl 1984a, VI §27, 624)<sup>22</sup>

Glauben, Wünschen und Zweifeln sind positive Modi. Gilt das auch für das blosse Dahingestelltseinlassen, wenn ich es zum Beispiel bloss dahingestellt sein lasse, ob *p*? Betrachten wir dazu die Beispiele, die Husserl für die Vorstellung als Akt mit „eine[r] eigenartige[n] Aktqualität“ gibt:

... jedes nichtausdrückliche ‚blosse Vorschwebenhaben‘ von Gedanken, ohne jedwede ‚Stellungnahme‘, ebenso jedes ‚blosse‘ Phantasieren usw. (Husserl 1984, V §32, 474)<sup>23</sup>

Nun ist das Phantasieren, ja auch das blosse Phantasieren, dass *p*, ein positiver Modus. Wie verhält es sich aber mit dem blossen Vorschweben desselben Gedankens? Ist es eine positive Aktqualität im selben Sinne? *Setzend* und *nicht-setzend* sind, könnte man meinen, Merkmale von positiven Qualitäten: Urteilen und einfaches Sehen haben das Merkmal [+setzend], visuelles Phantasieren und Annehmen haben das Merkmal [-setzend]. Ist aber das blosse Vorschweben eines Gedankens nicht die Abwesenheit jeglicher Aktqualität? Bei der Untersuchung von Husserls Analyse der Phantasie werden wir immer wieder auf Varianten dieser Frage stossen.

---

22. Vgl. (Husserl 1984, V §20, 429).

23. Vgl. auch a. a. O., 472.

Vergleichen wir nun die Auffassungen der nicht-ernsthafte Akte bei Meinong und Husserl und betrachten wir dabei zuerst die objektivierenden Akte. Man könnte meinen, Husserl verteidige 1901 eine Variante von MT1, der 50:50 These, die wir Meinong zugeschrieben haben, nämlich:

HT1

Zu jedem setzenden Acte gehört ein möglicher nichtsetzender Act von derselben Materie und umgekehrt. (Husserl 1984, V §34, 483)

Husserl schreibt auch:

Für die Zusammenfassung der nominalen und propositionalen Acte in Eine Klasse fiel für uns der Umstand entscheidend ins Gewicht, dass diese ganze Klasse durch einen qualitativen Gegensatz charakterisirt war, dass also, wie zu jedem nominalen *belief*, so zu jedem propositionalen, zu jedem vollen Urtheil, eine blosser Vorstellung als ihr Gegenstück gehöre. (Husserl 1984, V §34, 505)

Wie Meinong ein Jahr später, verwendet Husserl hier das Beispiel des Lesens eines Romans:

Die Urtheile werden zwar in gewisser Weise vollzogen, aber sie haben nicht den Charakter wirklicher Urtheile; wir glauben nicht, wir leugnen und bezweifeln aber auch nicht, was da erzählt wird; ohne jedes Fürwahrhalten lassen wir es auf uns wirken, wir vollziehen statt der wirklichen Urtheile bloss Einbildungen. Nicht als ob die Urtheile nun zu Gegenständen von Einbildungen würden. Wir vollziehen vielmehr statt des Urtheils als der „Fürwahrhaltung“ seines Sachverhalts, eine „Einbildung“ genau desselben Sachverhalts. (Husserl 1984, V §40, 510)

1913 modifiziert Husserl diese Stelle folgendermassen:

... wir vollziehen statt der wirklichen Urtheile bloss „Einbildungen“. Aber nun dürfte die Rede nicht etwa so verstanden werden, wie sie es gerade nahelegt, *als ob an die Stelle der wirklichen Urtheile Phantasieurtheile* treten sollten. Wir vollziehen vielmehr statt des Urtheils als der „Fürwahrhaltung“ seines Sachverhalts die *qualitative* Modifikation, das neutrale Dahingestellthaben desselben Sachverhalts, das keineswegs mit einem Phantasieren dieses Sachverhalts identifiziert werden darf. (Husserl 1984, V §40, 510)

Den Akten des Romanlesers wird hier die Qualität des neutralen Dahinstellens zugeschrieben, die nicht die Qualität eines Phantasieurtheils

– Meinongs Terminus – ist, und auch nicht die Qualität irgendeines Phantasierens. Daher erneut die Frage: Ist das neutrale Dahinstellen wirklich eine Aktqualität?

#### §4 *Austrian Entertainment*

Bekanntlich waren Bolzano und Frege der Meinung, Personen können in einer merkwürdigen Beziehung zu Abstrakta wie Sätzen an sich, Vorstellungen an sich und Gedanken stehen, nämlich in der Beziehung des *Fassens* oder des *Erfassens*. Philosophen aus Cambridge wie z. B. C. D. Broad und R. Braithwaite haben daraufhin ganz ähnlich gemeint, Personen können in der Beziehung des „entertainment“ zu „propositions“ stehen. Die folgenden Sätze sind inakzeptabel oder widersinnig:

\*Sam fasste, dass Maria traurig ist.

\*Sam entertained that Maria is sad.

(Man kann vielleicht sagen: „Sam erfasste, dass Maria traurig war“. Aber dann hat „erfasste“ nicht die Bedeutung, die sie in der „entertainment“ Tradition haben sollte). Das deutet darauf hin, dass wir mit diesen Verben keine Modi oder Aktqualitäten zuschreiben können, schliesst aber natürlich nicht aus, dass es ein Fassen von Gedanken gibt. Hat ein solches Fassen die Eigenschaft der Intentionalität? Husserl zufolge gibt es keine Intentionalität ohne Aktqualitäten.

Husserls Philosophie der Bedeutungen und unserer Beziehungen zu Bedeutungen in den *Logischen Untersuchungen* scheint gänzlich ohne eine Beziehung des Fassens oder Erfassens zwischen Personen und Bedeutungen auskommen zu *können*. Husserls Alternative zur „entertainment“ Tradition ist die These, dass das Meinen oder Bedeuten, das Meinen eines Gegenstandes und/oder das Meinen, dass *p*, immer *Einzelfälle von Bedeutungen* sind, wobei die Beziehung *ist ein Einzelfall von* keine intentionale Beziehung und das Meinen oder das Bedeuten keine Aktqualität ist. Urteilen und Begehren hingegen sind Aktqualitäten. Wer urteilt, dass *p* oder begehrt, dass *p*, meint – in dem Sinne, den Husserl mit dem Wort verbindet –, dass *p*<sup>24</sup>. Das Meinen, von dem Husserl spricht, ist weder eine Art Glauben noch eine Art Beabsichtigen.

---

24. Vgl. dazu (Mulligan 2011).

Obwohl Husserls erste Philosophie dem Fassen oder „entertainment“ eines propositionalen Gehalts anscheinend keinen Platz einräumt, vermute ich, dass die Anziehungskraft dieser Auffassung hinter Husserls wiederholten Hinweisen auf ein neutrales Dahinstellen steht, das zwar eine Aktqualität sein soll, sich aber weder mittels irgendwelcher existierenden Verben oder Adjektive belegen noch sonst positiv umschreiben lässt. Es ist wohl kein Zufall, dass man das Verhalten von jemandem, der einen Gedanken oder eine Satzbedeutung ohne zu urteilen oder zu glauben fasst, erfasst oder „entertains“, sehr natürlich als ein neutrales Verhalten beschreiben kann.

In der „entertainment“-Tradition erwähnt man oft den Fall des blossen Verstehens eines Satzes als Beleg für die Existenz von „entertainment“. Auch Husserl diskutiert das bloße Verstehen eines Ausdrucks. Gibt es „einen eigenen Akt des Verständnisses“?

Bei genauerer Betrachtung, zeigt es sich vielmehr, dass sich dieses Erlebnis [das „blosse Verständnis“ des Aussagesatzes] zum aktuellen Urteil analog verhält, wie etwa die blosse Phantasievorstellung zur genau entsprechenden Wahrnehmung. (Husserl 1984, V §30, 470)<sup>25</sup>

Auch beim bloßen Verstehen soll hier also eine Qualität im Spiel sein: „*isoliert kann die Materie nicht vorkommen*“ (Husserl 1984, V §31, 472; Hervorhebung K. M.). Doch welche Qualität, wenn nicht die Aktqualität des Phantasieurteilens? Soweit ich sehen kann, gibt Husserl keine klare und eindeutige Antwort auf diese Frage. Allerdings führt er im Laufe seiner Diskussion des Verstehens sehr früh eine Redewendung ein (als Teil einer zu kritisierenden These), die für ihn immer wichtiger wird:

Dabei befasste der Titel Vorstellen das ‚blosse‘ Vorstellen, eine Aktart, die wir uns durch das blosse Verständnis herausgerissener Worte u. dgl. exemplarisch verdeutlichen, oder auch durch das blosse Verständnis von gehörten Aussagesätzen, zu denen wir uns selbst völlig neutral verhalten. (Husserl 1984, V §33, 477)

Die Neutralität, von der hier bereits 1901 die Rede ist, scheint keine gute Charakterisierung einer positiven Aktqualität zu sein. Dasselbe gilt für die von Husserl mit Vorliebe verwendete Familie von Ausdrücken wie „Dahingestellthaben“ oder „Dahinstehendhaben“ (Husserl 1984, V §31, 472). Aber Husserl scheint auch der Meinung zu sein, das Dahingestellthaben, ob *p* sei ein Akt und müsse deshalb eine Aktqualität haben.

25. 1913 schreibt Husserl statt „Wahrnehmung“ „Erinnerung“.

In einem Brief an Meinong schreibt Husserl 1902 dann ausdrücklich, dass das Annehmen, dass  $p$  – nach Meinong ein Phantasieurteil – nicht mit dem Verstehen eines Satzes zu verwechseln sei:

Vollbewusst trenne ich die Assumptionen, die eigentlichen Annahmen, u. desgl. die *unter* Assumption gefällten belief-Acte von den bloss propositionalen Vorstellungen (nicht Vorstellungen *von* Urteilen), den Fällen schlichten Satzverständnisses – da ich hier einen descriptiven Unterschied zu finden glaube. (Meinong 1965, 106)<sup>26</sup>

Merkwürdigerweise überlegt sich Meinong 1910 in der zweiten Ausgabe von *Über Annahmen*, ob man nicht das Annehmen, dass  $p$  als einen Grenzfall des Urteilens, dass  $p$  betrachten soll:

Natürlich wird man ... ohne Zweifel den Unterschied zwischen Urteil und Annahme zunächst auf die Aktqualität beziehen dürfen; es wäre nichts Auffallendes, wenn eine nähere Beschreibung dieses Unterschiedes untunlich sein sollte. Hier eröffnet indes der Umstand, dass das Urteil sich in so natürlicher Weise als der gewissermassen reichere Tatbestand darstellt gegenüber der Annahme als dem ärmeren, die Aussicht, noch einen Schritt weiter zu gelangen. Es hat sich uns wiederholt als völlig natürlich ergeben, vorerst ohne jeden Anspruch auf Exaktheit zu sagen: in der Annahme haben wir im Grunde das Nämliche vor uns wie im Urteil, nur sozusagen um das Überzeugungs- oder Glaubensmoment weniger. Man könnte jetzt immerhin versuchen, das genau zu meinen in der Position: *die Annahme ist eine Art Grenzfall des Urteiles, charakterisiert durch den Nullwert der Überzeugungsstärke*. Es würde darin zugleich die Behauptung liegen, *Urteil und Annahme stimmen* im übrigen nicht nur dem Inhalte, sondern *auch dem Akte nach überein*, was aber natürlich nicht selbstverständlich ist, so dass es noch durchaus eines näheren Erweises bedürftig bliebe. (Meinong 1977, §59, 344; Hervorhebung K. M.)

Eine Konsequenz dieses Vorschlags ist wahrscheinlich, dass das Annehmen nicht mehr als eine Art des Phantasierens, nämlich als ein Phantasieurteil, betrachtet werden kann.

Die in Husserls *Ideen* (Husserl 1950, §112, 270) diskutierte „Neutralitätsmodifikation“ entspricht der qualitativen Modifikation, die Husserl in

---

26. Was sind die unter Annahmen oder Assumptionen gefällten „belief-Acte“? Es handelt sich nicht um Urteile mit einem konditionalen Inhalt – Sam urteilt, dass (wenn  $p$ , dann  $q$ ) – sondern um Urteile unter einer Annahme – unter der Annahme, dass  $p$ , urteilt Sam, dass  $q$ . Wenig später wird Reinach zwischen einem Versprechen mit einem konditionalen Inhalt und einem konditionalen Versprechen unterscheiden. Zur Kategorie „der unter Assumption gefällten belief-Acte“ vgl. (Edgington 1995, Mackie 1973, 92–108).

den *Logischen Untersuchungen* einführt. Um sie zu umschreiben, wiederholt und multipliziert Husserl die Metaphern:

Es handelt sich um eine Modifikation, die jede doxische Modalität, auf die sie bezogen wird, in gewisser Weise völlig aufhebt, völlig entkräftet...[S]ie ist das bewusstseinsmässige Gegenstück alles Leistens: dessen *Neutralisierung*. Sie liegt beschlossen in jedem sich des Leistens enthalten, es „ausser Aktion“ setzen, es „einklammern“, „dahingestellt sein lassen“ und nun „dahingestellt“ haben, sich in das Leisten „hineindenken“, bzw. das Geleistete „bloss denken“, ohne „mitzutun“ ... Der Setzungscharakter ist kraftlos geworden. Der Glaube ist nun ernstlich kein Glaube mehr, das Vermuten nicht ernstlich Vermuten, das Negieren nicht ernstlich Negieren usw. Es ist „neutralisiertes“ Glauben, Vermuten, Negieren u. dgl. ... (Husserl 1950, §109, 265f.)

Jemand, der diese Stelle (und ihresgleichen) gelesen hat, kann zweifelsohne sofort verstehen, was Edith Stein im Sinne hatte, als sie jemanden als „eine Neutralitätsmodifikation eines Menschen“ charakterisierte. Kann man aber die Aussage dieser Stelle selbst verstehen? Fragt man sich nicht erneut, was für eine Aktqualität denn das *neutralisierte Glauben*, dass *p* sein soll?

Keine gute Antwort auf diese Frage wäre „bloss Denken“, denn nach Husserls eigener Auffassung ist das Meinen eine Unterart des Denkens und, wie wir gesehen haben, keine Aktqualität. Auch als Umschreibung für die Neutralitätsmodifikationen im Allgemeinen ist „bloss Denken“ denkbar schlecht geeignet. Eine der Errungenschaften der *Untersuchungen* ist ja der Erweis der These, dass einfaches Sehen eines Dinges kein Urteilen sei, kein Meinen enthalte und keinen begrifflichen Inhalt habe, also kein prädikatives Denken sei.<sup>27</sup> Warum sollten wir unter diesen Voraussetzungen also denken, dass neutralisiertes Sehen, oder „Visualisieren“, ein Denken sei?

Wer behauptet, jemand lasse es dahingestellt, ob *p*, schreibt dieser Person ein intellektuelles Verhalten zu. „Entertain“, „fassen“ und „erfassen“ sind Ausdrücke, die einer Person wohl auch eine intellektuelle Beziehung zu Gedanken zuschreiben sollen. Husserl scheint aber der Meinung zu sein, es gäbe auch eine Neutralitätsmodifikation von Gefühlen und Begehungen. Die Neutralitätsmodifikation sei eine „gar nicht speziell zur Glaubenssphäre gehörige“.

Vielleicht kann uns ein Vorschlag von Edith Stein hier weiterhelfen. Wenn ich glaube, dass *p*, kann ich, so behauptet sie, „mich ganz so ver-

27. Vgl. dazu (Mulligan 1995).

halten, als wäre [dieser Glaube] nicht vorhanden, ich kann ihn unwirksam machen“ (Stein 1970, 43). Von diesem Verhalten behauptet sie als ausgezeichnete Kennerin von Husserls Philosophie, es sei „das Verhalten, das Husserl als epoché bezeichnet hat, die unwirksam gemachten Akte sind ‚neutralisierte‘“. Sie gibt dazu folgendes Beispiel:

Ich erwarte ... eine Nachricht, die mich veranlassen wird, eine Reise zu machen. Ich höre dann von nicht zuständiger Seite, daß das betreffende Ereignis eingetreten ist, und der Glaube daran stellt sich ohne weiteres ein. Aber ich „will“ nicht glauben, solange ich keine verbürgte Nachricht habe. Ich verhalte mich so, wie ich es tun würde, wenn der Glaube nicht vorhanden wäre: ich treffe keine Vorbereitungen, ich gehe meinem gewöhnlichen Tagewerk nach usw. – trotzdem die Glaubensstellungnahme unleugbar da ist. (Stein 1970, 43)

Das beschriebene Verhalten nennt Stein ein Ablehnen, sein Gegenteil das Annehmen. Ablehnen und Annehmen, wie sie hier aufgefasst werden, kommen in zwei Formen vor. Sie kommen als Akte vor, deren Gegenstände ausschliesslich die eigenen Stellungnahmen sind. Beispiele für Stellungnahmen sind der Glaube oder die Überzeugung, dass  $p$ , der Glaube an etwas oder jemanden, und viele Gefühle. Stellungnahmen sind entweder propositional oder nicht-propositional (der Glaube an etwas, die Bewunderung). Stellungnahmen setzen „Kenntnisnahmen“ – das Erkennen, dass  $p$ , das Erfassen eines Sachverhaltes, die Wahrnehmung eines Dinges – oder vermeintliche Kenntnisnahmen voraus. Wer seine Überzeugung annimmt, nimmt sie auf, bekennt sich zu ihr, stellt sich auf ihren Boden, lässt sie wirksam werden. Wer sie ablehnt, macht sie unwirksam. Ablehnen und Annehmen sind „freie Akte“, Stellungnahmen hingegen „ergreifen von mir Besitz“. Wer seine Stellungnahme annimmt oder ablehnt, nimmt zwar auch Stellung zu ihr, aber „in einem neuen Sinne“. Solche Akte haben eine eigenartige Eigenschaft: „Sie können nicht, etwa erst im Hintergrund und sich leise regend, allmählich von mir Besitz ergreifen“ – wie Gefühle und Überzeugungen –, „sondern ich muß sie aus mir heraus erzeugen, gleichsam geistig einen Streich führen“. Ablehnen und Annehmen haben, wie Stellungnahmen, Motive und Gründe, die zusammenfallen können oder auch nicht.<sup>28</sup>

---

28. Steins Auffassung des Annehmens unterscheidet sich in vielen Hinsichten von Jonathan L. Cohens (1992) einflussreicher Theorie der „acceptance“. Cohen unterscheidet zwischen der „acceptance“ oder Anerkennung, dass  $p$  und dem Glauben, dass  $p$ . Nach Stein ist der Gegenstand des Annehmens immer eine eigene Stellungnahme, die propositional oder nicht-propositional

Ablehnen und Annehmen treten aber nicht immer als eigenständige Akte auf:

Diese Annahme oder Ablehnung einer Stellungnahme, denen sie den Charakter der vollen Lebendigkeit und Wirksamkeit oder der Neutralität verdankt, brauchen keine selbstständig vollzogenen Akte zu sein, sie kann von vornherein mit dem einen oder anderen Charakter ausgestattet auftreten. Sie können aber auch jederzeit als eigene Akte vollzogen werden. (Stein 1970, 44)<sup>29</sup>

Annehmen und Ablehnen scheinen zur selben Kategorie von *Aktcharakteren* zu gehören, die Husserl im Sinn hat, wenn er sagt, dass eine Überzeugung oder eine Bewunderung sich mit unterschiedlichem Gewissheitscharakter einstellen kann: so gibt es eine gewisse, aber auch eine zweifelnde Überzeugung, und eine Bewunderung kann ungewiss sein. Ganz ähnlich scheint Stein der Meinung zu sein, dass man auf ablehnende Art überzeugt sein kann, auf annehmende Art bewundern kann, usw.

Wie verhält sich nun das Ablehnen à la Stein zur Neutralitätsmodifikation, die Husserl in §109 seiner *Ideen* einführt? Wenn es sich um dasselbe Phänomen handelt, dann hat Stein ihren Finger auf die Aktqualität gelegt, die dem Neutralisieren wesentlich ist – die Aktqualität des Ablehnens in ihrem Sinne. Husserl dagegen behauptet, es fehle „in der allgemeinen Sprache an einem eindeutigen Namen“ für die Neutralitätsmodifikation. Nach Stein erfolgt die Neutralisierung, indem ich entweder eine schon vorhandene Stellungnahme ablehne, oder indem eine Stellungnahme sich bereits mit dem Charakter der Ablehnung einstellt.<sup>30</sup> Husserl scheint die Neutralisierung als die Entkräftung eines vorhandenen Aktes aufzufassen. Er sagt, das Ergebnis der Neutralisierung des Vermutens sei, dass das Vermuten nicht „ernstlich“ sei. Wenn Stein recht hat, dann hat die Neutralisierung aber anscheinend wenig mit dem Phantasieren zu tun. Nach Steins Interpretation soll man nicht-ernstliches Vermuten nicht als ein Als-ob-Vermuten, sondern vielmehr als eine Ablehnung seines Vermutens, ein Nicht-zu-seinem-Vermuten-Stehen, auffassen. Schliesslich gibt Stein noch eine klare Antwort auf die Frage, was es heisst, nicht nur intellektu-

---

sein kann. Der Zusammenhang zwischen den zwei Auffassungen ist vielleicht dieser: wenn  $x$  seinen Glauben, dass  $p$  à la Stein annimmt, dann akzeptiert  $x$ , dass  $p$  à la Cohen. Beide Philosophen verwenden ihre Unterscheidungen zwischen Annehmen und Glauben in sehr eleganter Weise, um das Phänomen des Selbstbetrugs und der Selbsttäuschung zu verstehen (Stein 1970, 45f., Cohen 1992 §§23-28; vgl. dazu Mulligan 2013).

29. Vgl. a. a. O., 47.

30. Einmal erwähnt Stein die Möglichkeit der Annahme einer nicht vorhandenen Stellungnahme (Stein 1970, 44).

elle sondern auch affektive Stellungnahmen zu neutralisieren. Andererseits schreibt Husserl ausdrücklich, man könne „alles Willentliche“ aus dem „Dahingestellt-sein-lassen“ der Neutralitätsmodifikation ausschalten. (Im „entertainment“ oder im Fassen eines Gedankens kommt auch nichts „Willentliches“ vor).

Das einzige von Husserl ausgearbeitete Beispiel der Neutralisierung passt kaum zu der von Stein vorgelegten Interpretation. Husserl beschreibt in §111 der *Ideen* die ästhetische Betrachtung von Dürers Kupferstich „Ritter, Tod, und Teufel“ als Beispiel des „neutralen Bildobjektbewusstseins“, das eine Neutralisierung der „normalen, in unmodifizierter Gewissheit setzenden Wahrnehmung“ sein soll. Er unterscheidet dabei die normale Wahrnehmung des Kupferstichblattes, dann „das perzeptive Bewusstsein, in dem uns in den schwarzen Linien farblose Figürchen ‚Ritter auf dem Pferde‘, ‚Tod‘ und ‚Teufel‘ erscheinen“, und drittens die im Bilde abgebildeten „Realitäten, dem Ritter aus Fleisch und Blut“. Das Bewusstsein des Bildobjekts ist, behauptet er, eine Neutralitätsmodifikation der Wahrnehmung. Das Bildobjekt ist als „gleichsam-seiend“ bewusst. Und auch das Bewusstsein der abgebildeten Realitäten soll eine Neutralisierung sein. Stein zufolge kann die Wahrnehmung jedoch nicht unwirksam gemacht oder abgelehnt werden, da sie eine Art Kenntnissnahme ist.<sup>31</sup> Trotzdem erinnert eine zusätzliche von Husserl eingeführte Unterscheidung an Steins spätere Unterscheidung zwischen Akten und Aktcharakteren: die Neutralisierung, die das Bewusstsein der abgebildeten Objekte ausmacht, darf nicht, so Husserl, „als eine an eine vorgängige Setzung angeschlossene umbildende Operation“ aufgefasst werden, selbst wenn sie das gelegentlich auch sein kann.

Auch in seiner Beschreibung des Beispiels des „neutralen Bildobjektbewusstseins“ weigert Husserl sich, die Aktqualität oder den Modus der Neutralisierung anzugeben. Er scheint die Existenz namenloser Modi zu vertreten wie Aristoteles die Existenz namenloser Tugenden. Auch wenn es für uns nun naheliegt, das Bewusstsein des abgebildeten Ritters als eine Art visuelles „entertainment“ oder Fassen zu beschreiben, gibt es dazu allerdings noch eine Alternative. *Im Sinne* Meinongs könnte man vielleicht sagen, dass, wer den Dürerschen Kupferstich ästhetisch betrachtet, das Blatt sieht und einen Ritter *sieht<sub>pb</sub>*. Je mehr das Als-ob-Sehen des Ritters vom Sehen des Blattes und seiner Eigenschaften gesteuert wird, desto eher ist es ästhetisch adäquat.

31. 1912 behauptet Husserl auch, im schlichten Wahrnehmen gebe es kein Stellungnehmen und die schlichte Wahrnehmung könne nicht „umgewandelt“ – oder wie er später sagen wird: neutralisiert – werden (Husserl 1980, Nr. 15j, 401, 403f.).

Husserl zufolge ist neutralisiertes Glauben, dass *p* jedoch kein Phantasieren. In der Tat reserviert Husserl in den *Ideen* das Wort „Phantasie“ für ein Phänomen, das er manchmal, aber nicht immer, als eine *Unterart* der Neutralitätsmodifikation charakterisiert:

... die Phantasie [ist] in der Tat eine Neutralitätsmodifikation, ... anwendbar auf alle Erlebnisse, ... [muss aber] von der allgemeinen Neutralitätsmodifikation ... unterschieden werden ... (Husserl 1950, §111, 267f.)

[D]as Phantasieren [ist] die Neutralitätsmodifikation der „setzenden“ Vergegenwärtigung, also der Erinnerung im denkbar weitesten Sinne (Husserl 1950, §111, 268)

Wer phantasiert, erinnert sich nicht, er {erinnert sich} sozusagen. Nach Husserl hat die Erinnerung „die merkwürdige Eigentümlichkeit, dass sie auch Vergegenwärtigung der Wahrnehmung ist, und nicht nur Vergegenwärtigung des Wahrgenommenen“ (Husserl 1980, 306f.). Wenn wir annehmen, dass der Gegenstand des Erinnerns ein Erlebnis *und* der Gegenstand dieses Erlebens ist, dann dürfen wir vielleicht behaupten:

Wenn *x* phantasiert, dann {erinnert sich} *x* {an} ein Sehen/Hören/... von *y* oder {an} ein Sehen, dass *p*...

Damit ergeben sich zwei Analogien zwischen Husserls Phantasieauffassung von 1913 und Martins Erläuterung der sinnlichen Phantasie. Erstens: Zu Husserls Phantasien gehört der von Martin beschriebene Fall, in dem der Gegenstand des Phantasierens ein sinnliches Erlebnis und auch dessen Gegenstand ist. Zweitens: alle Arten des Phantasierens à la Husserl unterscheiden sich wie alle Arten des sinnlichen Phantasierens à la Martin nur hinsichtlich ihrer Gegenstände (oder Materien oder Inhalte). Das Phantasieren ist immer dieselbe spezifische Aktqualität und lässt keine weitere Differenzierung zu. Es ist immer von derselben Art. Andererseits ist das husserlsche Phantasieren im Gegensatz zum martinschen Phantasieren eine Modifikation des Erinnerns (im weitesten Sinne des Wortes). Husserl behauptet auch,

die Phantasie-modifikation als Vergegenwärtigung [ist] iterierbar, während die Wiederholung der „Operation“ der Neutralisierung wesensmässig ausgeschlossen ist. (Husserl 1950, §112, 270)

(Hier scheint Husserl das Phantasieren *nicht* als eine Art Neutralisierung zu betrachten.) Das heisst,

Wenn  $x$  {sich erinnert}, dann kann  $x$  später {sich erinnern an} sein {sich Erinnern}<sup>32</sup>

Wenn Steins Interpretation der Neutralisierung zutrifft, kann das, was sie „Ablehnung“ oder „unwirksam machen“ nennt, auch nicht iterierbar sein. Dem ist auch so, denn nur Stellungnahmen können abgelehnt werden. Und wie wir gesehen haben, ist das Ablehnen (wie auch das Annehmen) keine Stellungnahme, sondern eine Operation, in der man „in einem neuen Sinne“ Stellung nimmt.

Das Phantasieren ist nicht die einzige Unterart der nicht-ernsten objektivierenden Akte in den *Ideen*. Husserl führt dort eine zweite Unterart, „eine neue Modifikation“, ein: das Annehmen oder Ansetzen:

Freilich kann jedes blosse Sich-denken [= Neutralitätsmodifikation] in ein Annehmen, Ansetzen übergeführt werden, ... Aber Ansetzen ist wiederum so etwas wie setzen, ... nur dass es eine ganz eigene, der oben behandelten Hauptreihe [= Neutralitätsmodifikation] gegenüber und seitab stehende Modifikation der Glaubenssetzung ist (Husserl 1950, §110, 267)<sup>33</sup>

So wie das Phantasieren (im Jahre 1913) eine Modifikation des Sich-Erinnerns sein soll, so soll das Annehmen eine Modifikation des Glaubens oder Urteilens sein. Es ist aber keine unmittelbare Modifikation davon. Es soll eine Modifikation der Neutralitätsmodifikation der Glaubenssetzung sein („übergeführt“). Husserl sagt von den „eigentlichen nichtneutralisierten Noesen“, von den „ernsten“ Akten (Meinong), dass sie

ihrem Wesen nach einer „Rechtsprechung der Vernunft“ unterliegen, während für die neutralisierten die Frage nach Vernunft und Unvernunft keinen Sinn gibt. ... Der „blosse Gedanke“ von Wirklichkeiten, Möglichkeiten usw. „prätendiert“ nichts, er ist weder als richtig anzuerkennen, noch als unrichtig zu verwerfen. (Husserl 1950, §110, 224)

Genau das soll von den Annahmen aber nicht gelten:

Ansetzen ... *kann* in die Einheit vernunftmässig zu beurteilender Setzungen als Glied (ihr Ansatz als hypothetischer „Vordersatz“ oder Nachsatz) eintreten und damit selbst der Vernunftwertung unterzogen werden. Nicht von einem bloss dahinstehenden Gedanken, aber wohl von einem hypothetischen Ansatz kann es heissen, er sei ein richtiger oder nicht. (Husserl 1950, §110, 224; Hervorhebung K. M.)

32. Das Phantasieren ist iterierbar wie schon die „vorstellende Objektivierung“ der *Untersuchungen* (Husserl 1984, V §39).

33. Vgl auch den schon zitierten Brief (1902) von Husserl an Meinong.

Mindestens drei Fragen drängen sich hier auf. Wie unterscheidet sich ein frei schwebendes Ansetzen, ein Ansetzen, das kein Glied ist, von einem bloss dahinstehenden Gedanken? Wenn die Aktqualität eines bloss dahinstehenden Gedankens weder das Annehmen noch das Ansetzen ist, was ist sie dann? Eine Antwort auf die erste Frage wäre, dass das Wort „kann“ in der zitierten Stelle ein Lapsus war; vielleicht meint Husserl eigentlich, es gehöre zum Wesen eines Ansetzens oder Annehmens, ein Glied zu sein.<sup>34</sup> Schliesslich kann man auch fragen, ob das Ansetzen wirklich eine Neutralisierung voraussetzt.

Dass das Annehmen ein Glied sein *kann* ist selbstverständlich. Es kann als Glied eines Schliessens vorkommen. Es kommt auch als Glied eines bedingten Urteilens vor: dann urteilt man, dass  $q$ , unter der Annahme, dass  $p$ . Mackie schreibt in diesem Zusammenhang von dem wichtigen aber „perhaps obscure concept of asserting something within the scope of a supposition“ (Mackie 1973, 92). Wenn Husserl und Meinong recht haben, gibt es hier aber keine Dunkelheit. Wer bedingt urteilt, verbindet ein Annehmen und ein Urteilen mit Hilfe einer „intentionalen Verknüpfung“ (Husserl 1950, §118, 291). Eine intentionale Verknüpfung ist eine geistige Operation (Meinong, Husserl) oder Funktion (Reinach, Pfänder). Dieser Fall ist von zwei anderen Fällen zu unterscheiden. Erstens, der Fall, in dem man urteilt, dass (wenn  $p$ , dann  $q$ ). Hier verwendet man den Begriff, der durch „wenn...dann“ ausgedrückt wird. Wer dagegen unter der Annahme, dass  $p$ , urteilt, dass  $q$ , verwendet den Begriff, der durch „wenn...dann“ ausgedrückt wird, nicht. Zweitens, der Fall des bedingten Versprechens (Befehlens). Jemand, der ein bedingtes Versprechen macht, indem er sagt:

Wenn  $p$ , dann verspreche ich Dir hiermit, dass  $q$ ,

verwendet wohl den Ausdruck „wenn ... dann“, bildet damit aber keinen wahrheitsfähigen Satz und urteilt auch nicht. Husserl und Meinong zufolge gibt es neben dem Urteilen unter einer Annahme noch andere geistige Operationen, wie zum Beispiel das Kolligieren und das Disjungieren. Und es kommt oft vor, dass man handelt, begehrt oder sich empört *auf Grund*

---

34. Vgl. (Künne 1995), der darauf hinweist, dass diese Idee bei (Heller 1929, 509ff.), (Waismann 1976, 436-9) und (Dummett 1973, 313) zu finden sei. Der bei Philosophen beliebte Ausdruck „Nehmen wir an, dass ...“ wie auch das Aufstellen einer Annahme in einem Argument haben sehr wohl immer einen einführenden Charakter und funktionieren als eine Art Prolog.

der Tatsache, dass *p*, ohne dabei den Begriff „weil“ zu verwenden.<sup>35</sup> *Muss* aber das Annehmen ein Glied sein? Die vom Romanleser vollzogenen quasi-Urteile, wie sie Ingarden in *Das literarische Kunstwerk* darstellt, sind keine Glieder der gerade erwähnten Arten.

Fragen wir uns nun: Wie verhält sich Husserls Auffassung der Neutralitätsmodifikation, der Phantasie und des Annehmens zu Meinongs These (MT2), der zufolge ein Phantasie-N kein N ist, aber einem N ähnelt? Vieles spricht dafür, dass Husserl folgende drei Thesen akzeptiert:

- HT2a Ein neutralisiertes Sehen (Hören, Urteilen...) ist kein Sehen (Hören, Urteilen...), ähnelt aber einem solchen.
- HT2b Ein Phantasieren ist ein {Sich-Erinnern} und ein {Sich-Erinnern} ist kein Sich-Erinnern, ähnelt aber einem solchen.
- HT2c Ein Annehmen ist keine Glaubenssetzung (und kein Urteilen), ähnelt aber einem solchen *und* einem neutralisierten Urteilen.<sup>36</sup>

Wir wissen bereits, dass das Phantasieren laut Husserl 1913 ein neutralisiertes Sich-Erinnern sein soll. Wie verhält sich also ein neutralisiertes F-en zum F-en? Die Neutralisierung soll „keine Privation, sondern eine Modifikation“ sein (Husserl 1950, §111, 270). Als solche ist sie immer ein positives Phänomen, eine Aktqualität oder ein Modus. Husserl sagt wenig darüber, wie „Modifikation“ im Ausdruck „Neutralitätsmodifikation“ oder im Ausdruck „qualitative Modifikation“ zu verstehen sei. Man darf vielleicht annehmen, dass „Modifikation“ hier so zu verstehen ist wie in den von Husserl diskutierten Beispielen der *Bedeutungsmodifikation* (Husserl 1984, IV, 329–334). Husserl meint, die Bedeutung [dass Sam traurig ist] sei eine Modifikation der Bedeutung [Sam ist traurig], genauso wie der Ausdruck „dass Sam traurig ist“ eine sprachliche Modifikation des Ausdruckes „Sam ist traurig“ ist. Die eine Bedeutung ist eine Satzbedeutung, die andere eine nominale Bedeutung. Keine Satzbedeutung ist eine nominale Bedeutung. Aber die zwei Arten von Bedeutung teilen einige Eigenschaften. Beide enthalten z. B. die Bedeutung oder den Begriff [trau-

---

35. Vgl. (Mulligan 2010).

36. Diese drei Arten von nicht-ersten Akten sind keineswegs die einzigen Arten die Husserl beschreibt. Hier sehe ich z. B. von der „imaginativen Modifikation“ (Husserl 1984, 509–14) und der Unterscheidung zwischen perzeptiver und reproduktiver Phantasie ab. Ich gehe auch nicht auf die vielfachen Anwendungen ein, die Husserl von seiner Taxonomie der nicht-ersten Akten macht – z. B. in der Metaphysik, in der Analyse des Bildbewusstseins und in seiner Lehre von singulären negativen existenzialen Urteilen.

rig]. Das neutralisierte Sehen (Hören, Erinnern, Urteilen....) ähnelt dem Sehen (Hören, Erinnern, Urteilen...), genauer: es ähnelt der Aktqualität des Sehens (Hörens, Erinnerns, Urteilens...).<sup>37</sup> Meinong zufolge kann Phantasieren vielerlei sei. Und Husserl zufolge kann die Neutralität auch vielerlei sein. Wenn aber Stein recht hat, dann ist die Neutralisierung immer Ablehnung, und damit einerlei. Variabel ist bloss, *was* man ablehnt oder unwirksam macht, also die verschiedenen Stellungnahmen, nicht aber das Ablehnen selbst.

Eine Eigenschaft, die Husserl der Neutralitätsmodifikation und Meinong dem Phantasieren zuschreibt, betrifft das Verhältnis dieser Phänomene zum Wollen. Husserl sagt von der Neutralitätsmodifikation im Allgemeinen, von der Modifikation des Sich-Denkens, wie auch vom Annehmen oder Ansetzen, sie unterlägen „der unbedingt freien Willkür“ (Husserl 1950, §110, 267). D. h., ich kann mich entscheiden, den Gedanken, dass *p*, bloss zu denken, und ich kann mich auch entscheiden, anzunehmen, dass *p*. Urteilen, Glauben, Bewundern und Wollen dagegen sollen dem Wollen nicht direkt unterstehen. Obwohl die Neutralitätsmodifikationen direkt dem Wollen unterworfen sind, kann ein Phantasieren „auftauchen“, ohne Gegenstand irgendeines Wollens oder Begehrens zu sein.

Wenn wir uns nun fragen, wie sich Husserls Auffassung der nicht-ernsten objektivierenden Akte zu Meinongs Vermutung MV1 verhält,

MV1 „[D]ie Urtheile [geben] in irgend einem Sinne eine Art Oberstufe zu den Annahmen als Unterstufe [ab]“ (Meinong 1902, 280)

dann lautet die Antwort: Sie ist damit unverträglich. Husserl zufolge sind Urteile nicht eine Art Oberstufe zu Annahmen, genauso wenig wie Annahmen eine Art Unterstufe zu Urteilen sind. Vielmehr kommt jeder propositional gegliederten Materie genau eine objektivierende Aktqualität zu. Wenn ein Urteilen, dass *p* ein Annehmen, dass *p* voraussetzen würde, dann kämen derselben propositionalen Materie zwei episodischen Aktqualitäten zu. Auch jeder nicht-propositional gegliederten Materie kommt genau eine objektivierende Aktqualität zu. Wittgenstein sagt, man könne nicht gleichzeitig den Eiffelturm sehen und ihn sich vorstellen. So denkt auch Husserl, obwohl er andererseits meint, dass das Urteilen im Vergleich zum Annehmen die fundamentalere Aktqualität sei. Ihm zufolge soll das Annehmen eine Modifikation des Urteilens sein, so wie

---

37. Vgl. (Husserl 2002, 125).

das nicht-propositionale einfache visuelle Phantasieren eine Modifikation des einfachen nicht-propositionalen Sehens sein soll.

Broad erkennt die Schwäche von Meinongs Oberstufenlehre, wenn er sagt:

I cannot help thinking that there are really three different attitudes towards a proposition and that Meinong confounds two of them under the name Annahme. These two I would distinguish as supposition and entertainment. It seems to me that entertainment clearly differs from supposition and is presupposed both by it and by judgment. When Meinong insists on the resemblance of Annahmen to judgment I think he has suppositions in mind ; when he says that every judgment presupposes a corresponding Annahme I think he has entertainment in mind. But entertainment as distinct from supposition does not seem to me to differ from having an idea. (Broad 1913, 92)<sup>38</sup>

Husserls Darstellung der *nicht-ernsthaften objektivierenden Aktqualitäten* in den *Untersuchungen* und in den *Ideen*, so wie wir sie gerade skizziert haben, macht kaum einen expliziten Gebrauch von Meinongs Vorschlag, wonach alle Gegenstücke der ernsthaften Aktqualitäten Arten des Phantasierens seien. Dies ist nicht zuletzt eine Folge des engen Sinnes, den Husserl dem Terminus „Phantasie“ in den *Ideen* beilegt. Das Bild von Husserls Auffassung ändert sich jedoch beträchtlich, wenn man seine Auffassung der *nicht-ernsthaften nicht-objektivierenden Aktqualitäten* heranzieht, sowie andere seiner Schriften.

## §5 *Phantasiegefühle, Scheingefühle, Phantasiebegehungen, Scheinbegehungen*

Am 5. April 1902 schreibt Husserl an Meinong:

Die von Ihnen vertretene Auffassung, welche bloss nominale Vorstellungen von Sachverhalten und „Scheinurteile“ unterscheidet, und den gleichen Unterschied dann *bei allen anderen Acten durchführt*, war zeitweise auch meine eigene trotz der immerfort empfundenen Schwierigkeiten. Zeugnis dafür giebt ein mir übriggebliebenes Conceptblatt vom Ende 1899, das ich beizulegen mir gestatte. (Meinong 1965, 106; Hervorhebung K. M.)<sup>39</sup>

Am 25. September 1906 notiert Husserl zu Meinongs *Über Annahmen*:

---

38. Weshalb Broad *entertainment* als „attitude“ charakterisiert, ist mir unverständlich.

39. Husserl spricht von einem Scheinurteil schon 1894; vgl. (Husserl 1979, 318).

Ich finde in dem Buch nur einen bedeutsamen Gedanken, den ich in meinen „Logischen Untersuchungen“ nicht ausgesprochen habe, obschon ich ihn während der Ausarbeitung hatte und erwog, aber dann doch nicht aufzunehmen wagte: *die Übertragung der Modifikation von Urteilen in „blosse Vorstellungen“, auf Wünsche und alle sonstigen Akte*. Ich besitze noch das datierte Blatt über diese Frage (1894), in dem ich genau Meinongs Position einnahm. Aber freilich, ich sah grosse Schwierigkeiten, die Meinong nicht sieht, und die hinderten mich, Folge zu geben. (Husserl 1956, 295f.; Hervorhebung K. M.)

Irgendwann sah Husserl diese Schwierigkeiten nicht mehr und folgte *de facto* seinem österreichischen Landsmann:

*Jedem [Bewusstsein] können wir sein spielerisches Analogon gegenüberstellen. Phantasie bedeutet also eine allgemeine Modifikation, die in sich selbst zurückweist auf ein unmodifiziertes Bewusstsein. Das unmodifizierte nennen wir positionales Bewusstsein, das Sein und die Seinsmodalitäten die in ihm dem Gegenstand zuerteilt sind, sind ernstliche, sind Positionen. So entspricht nicht nur dem Erfahren als positionalem anschaulichem Bewusstsein individueller Gegenstände ein Phantasieerfahren, sondern auch dem wirklichen Wünschen ein Phantasiewünschen, Sich-in-ein-Wünschen-Hineinphantasieren, im Spiel, dem wirklichen Begehren und Wollen ein Phantasiebegehren etc. Die Korrelate der Phantasie sind Fiktionen, „Phantasiebilder“, die Korrelate der Positionalität sind Wirklichkeiten, vermeinte oder wahre Wirklichkeiten: wirklicher Gegenstand, wirklicher Trug, wirkliche Wahrscheinlichkeiten usw. In der Phantasie ist all das mit dem Vorzeichen des Spiels vermeint, Nichtsein wie Sein. (Husserl 2000, §3, 13; Hervorhebung K. M.)<sup>40</sup>*

Hier hat „Phantasieren“ denselben breiten Umfang wie bei Meinong. Und „Sich-in-ein-Wünschen-Hineinphantasieren“ lässt sich auf Englisch gut erläutern mit „imagining-wishing“ (vgl. (T1-2) oben). Hier ist das Phantasieren also nicht als das Ergebnis einer Operation mit einem vorhandenen Akte zu verstehen: wer ein Phantasiebegehren hat oder erlebt, braucht nicht zuerst begehrt zu haben.

Husserl spricht nun auch, wie Meinong, von „quasi-Gefühlen“:

Wir haben nicht nur die Gefühle des ästhetischen Wertens, sondern auch die in uns als „Reaktion“ geweckten Gefühle (oder *quasi*-Gefühle), Furcht und Mitleid etc., die durch die Erscheinungsweise und die durch die-

---

40. Vgl. (Husserl 1980, 339f.), (Husserl 1950, §117, 288), (Husserl 2002, 212f.), (Cairns 1973, 259–262).

selbe zunächst bestimmten Gefühle mitbeeinflusst werden. (Husserl 1980, Nr. 15, 389)<sup>41</sup>

Er ist auch bereit, von einem geistigen oder psychologischen Fingieren zu reden:

Ich kann mich nun auch als mich verändernden in sehr vielfacher Weise vorstellen und mich als anders seiend vorstellen. Aber mich als anders seiend fingieren, das heisst nicht, ein *anderes Subjekt*, ein zweites Subjekt, einen ‚Anderen‘ *mir gegenüber* fingieren. Ich bin noch immer, auch in der Fiktion, Ich. (Husserl 1973, 289)<sup>42</sup>

1937 beschreibt Husserl „das Reich der Phantasieakte“, wie Meinong, als das Reich „der Akte im Modus des Als-ob“:

Diese Modifikation des Als-ob steht allen wirklichen Akten, den „positionalen“ ... gegenüber. Jeder ... hat sein Gegenstück als mögliches Als-ob, und ich, das Vollzugsich dieser Akte im Modus des Als-ob, bin in ihnen allen bezogen nicht auf die wirkliche Welt, ... sondern ich bin phantasierendes Ich, Phantasien vollziehend, in ihnen in Quasiakten mich auf eine Phantasiewelt, eine mir im Als-ob geltende, beziehend.

In den *Ideen* habe ich den, wie ich späterhin sah, wenig passenden Terminus „Neutralitätsmodifikation“ für diese Als-ob-Abwandlungen von Akten gewählt. (Husserl 1993, 424f.)

Die letzte Bemerkung ist nicht ganz zutreffend. In den *Ideen* hatte Husserl es vermieden, alle Aktqualitäten, die zu den Neutralitätsmodifikationen gehören, als Als-ob-Abwandlungen der Phantasie zu charakterisieren, und hatte auch sorgfältig zwischen der allgemeinen Neutralitätsmodifikation und der Phantasie-modifikation unterschieden. Was er hier als Phantasieren beschreibt, scheint nichts anderes als das Phantasieren à la Meinong zu sein. Sehr spät scheint Husserl Meinong nun gänzlich recht zu geben.

Die Geschichte ist aber noch komplizierter. Innerhalb der Grazer Schule gibt es eine merkwürdige Entwicklung in Form eines Vorschlags von Ernst Mally (1912), der von Meinong aufgegriffen und ausgebaut wurde (1916). Mit diesem Vorschlag wird in einem gewissen Sinne Husserls Unterscheidung zwischen dem Phantasieren und den Neutralitätsmodifikationen ernst genommen und sogar verbessert. Es handelt sich um Meinongs

---

41. Vgl. a. a. O., 237f., 390, 520–24.

42. Der Fall den Husserl hier beschreibt – ich stelle mir mich als anders seiend vor – entspricht dem dritten Disjunkt in (T1) oben.

Unterscheidung zwischen *schattenhaften* und *urteilsartigen Annahmen* und um die Verallgemeinerung dieser Unterscheidung, die Meinong dazu führt, zwischen *schattenhaften Phantasiegefühlen* und *ernstartigen Phantasiegefühlen* zu unterscheiden. Damit erreicht die barocke Unterscheidungskunst der österreichischen Philosophie einen ihrer Höhepunkte.

Meinong bemerkt: obwohl es „eine weitgehende Äquivalenz“ zwischen Ausdrücken der Form „A existiert“ und solchen der Form „die Existenz des A“ gibt, ist „Es ist wahr, dass A existiert“ akzeptabel, nicht aber „Die Existenz von A ist wahr“. Er zitiert weiter den Vorschlag von (Mally 1912, 62):

E. Mally versucht dies nun im Hinblick darauf zu verstehen, daß es geschehen kann, daß wir ein Objektiv nicht *bloß denken*, sondern in der Annahme, „daß A ist“, es, wie man kurz sagen könnte, „phantasiemäßig urteilen“, d. h. indem wir in dieser Annahme das urteilsmäßige Erfassen des Objektivs nicht bloß dem Inhalte nach reproduzieren, sondern auch dem Urteilsakte, *dem Überzeugungsmomente* nach sozusagen nachahmen oder phantasiemäßig nachbilden, „so tun“, als urteilten wir es ... Neben solchen Annahmen, die sich als Phantasieurteile darstellen, gibt es aber allem Anscheine nach auch Annahmen, die bloß den Inhalt eines (möglichen) Urteiles wieder vollziehen, die sich als „*bloßes Denken*“ des Objektivs darbieten, als Setzung ohne Nachbildung des Überzeugungsmomentes ... (Meinong 1968, 332f.; Hervorhebung K. M.)

Es ist nicht klar, weshalb Mally meint, man könne phantasiemässig urteilen mit einem nicht-propositionalen Inhalt („dass A ist“ ist ein Name). Interessant ist jedoch der Vorschlag, zu einer Art Annahme gehöre ein *Als-ob*-Überzeugungsmoment, eine Überzeugung<sub>pb</sub>. Wie wir gesehen haben, charakterisierte Meinong 1902 die Annahme als etwas Urteilsartiges *ohne* irgendeine Überzeugung und 1910 lautete seine Vermutung, dass Annahmen eine Art Grenzfall des Urteilens sei, charakterisiert durch den Nullwert der Überzeugungsstärke. Von einer *Als-ob*-Überzeugung war keine Rede. Die zweite Art Annahme, die Mally einführt, wird von ihm als ein „*bloßes Denken*“ eines Objektivs charakterisiert. Diese Charakterisierung ist die von Husserl vorgezogene Charakterisierung der qualitativen Modifikation (1901) und der Neutralitätsmodifikation (1913), genau die Charakterisierung also, der wir keinen positiven Sinn geben konnten ohne die Idee des „*entertainment*“ heranzuziehen.

Meinong kommentiert Mallys Vorschlag wie folgt:

Ist nun ... Wahrheit (im Gegensatz zu Tatsächlichkeit) zunächst Sache der Erfassungsobjektive, dann liegt es in der Tat nahe, die durch Annahmen von

der ersten Art erfaßten Objektive eher wahr (resp. falsch) zu nennen als die durch Annahmen der zweiten Art erfaßten. (Meinong 1968, 333)

Wie wir gesehen haben, meint auch Husserl, dass ein bloßer Gedanke nichts „prätendiert“, dass er weder als richtig anzuerkennen noch als unrichtig zu verwerfen ist. Das Ansetzen oder Annehmen, dass  $p$ , als Glied z. B. eines konditionalen Urteilens, ist dagegen richtig oder unrichtig.

Meinong baut anschließend Mallys Vorschlag weiter aus:

Urteile lassen sich betrachten als Annahmen, zu denen das Glaubensmoment (in irgendeinem seiner Stärkegrade) hinzugetreten ist. Es kann nun geschehen, daß auch schon die Annahme ein Moment aufweist, das zwar noch kein Glaube, keine Überzeugtheit ist, so daß die Annahme darum noch nicht aufhört, Annahme zu sein, – das aber *glaubensähnlich* genug ist, um die Annahme als urteilsähnlich erscheinen zu lassen, was unter anderen Umständen des Annehmens, beim ‚bloßen Denken‘ noch nicht der Fall ist ... Die Bezeichnung „Phantasieurteil“ für die urteilsähnlichen Annahmen würde sich kaum empfehlen ... Dagegen wäre die Bezeichnung „*urteilsartige Annahme*“ nicht uncharakteristisch, die Bezeichnung „*schattenhafte Annahme*“ dagegen von der ersten Benennung ausreichend unabhängig gebildet, um vorerst in suspenso zu lassen, ob zwischen diesen beiden Typen des Annehmens etwa noch ein Mittleres zu konstatieren sein mag. (Meinong 1968, 333f.; Hervorhebung K. M.)<sup>43</sup>

Die urteilsartige Annahme lässt sich vielleicht so charakterisieren:

$x$  urteilt  $_{pb}$ , dass  $p$  &  $x$  glaubt  $_{pb}$ , dass  $p$

die schattenhafte Annahme dagegen als

$x$  urteilt  $_{pb}$ , dass  $p$  & Nicht ( $x$  glaubt  $_{pb}$ , dass  $p$ )

Es ist (mir) aber nicht klar, ob Meinong 1916 tatsächlich eine schattenhafte Annahme als eine Art Phantasieren charakterisieren will. Findlay gibt ein interessantes Beispiel für die Unterscheidung:

The need for these two varieties of assumption [urteilsartige Annahmen, schattenhafte Annahmen] is easily shown if we consider that it is quite easy to live through the transition from belief to disbelief *in a purely imagina-*

---

43. Zur parallelen Unterscheidung zwischen *schattenhaften Phantasiegefühlen* und *ernstartigen Phantasiegefühlen* vgl. (Meinong 1968, 334f.).

*tive way.* If I first put myself in the position of a person who believes some religious dogma which I do not believe, and then experience imaginatively the loss of belief in this dogma, I am living through an experience in which something resembling the moment of seriousness is at first present, and then disappears. Yet the whole experience is not a serious one at all. We can only understand this highly curious state of affairs if we hold that even an assumption may have a *simulacrum* of genuine conviction; we pass from an assumption which is judgment-like to one that is altogether shadowy. (Findlay 1963, 229f.)

## §6 *Schluss*

Die These, jedem ernstesten Akte entspreche sein phantasiemässiges Gegenstück, wobei jedes derartige Gegenstück eine angebbare Aktqualität besitze (vgl. T1–2 in §2), hat *prima facie* gewisse Vorteile gegenüber anderen Auffassungen der Phantasie.

Alan Whites (1990, Kap. 20) vorsichtige und gründliche Verteidigung der These, das Phantasieren („imagining“) bestehe darin, an Möglichkeiten zu denken, leidet daran, dass die Theorie entweder keinen Platz für das sinnliche Phantasieren hat oder –sehr unplausibel– das sinnliche Phantasieren als eine Art Denken auffasst. Martins Auffassung der sinnlichen Phantasie ist *eo ipso* keine Theorie der Phantasie *tout court*. Falls sie sich verallgemeinern lässt, müsste man wählen zwischen dem Ergebnis der Verallgemeinerung und Husserls Theorie der allgemeinen Neutralitätsmodifikation. Die österreichische Theorie der Phantasie, wovon einige Bruchstücke hier isoliert und erläutert wurden, hat – im Gegensatz zu neueren Versuchen, die Phantasie über „make-believe“ oder „simulation“ zu erfassen<sup>44</sup> – den Vorteil, dass sie sich mühelos in eine ausgearbeitete allgemeine Theorie der Intentionalität einfügt. Sie vermeidet den alten Fehler, das ästhetische Erleben als eine Art Täuschung oder Illusion aufzufassen. Und im Gegensatz zu kantischen und neo-kantischen<sup>45</sup> Auffassungen vermeidet es diese Theorie, das Phantasieren und die Einbildungskraft zu Bestandteilen oder Aspekten aller geistigen Episoden und Zustände zu machen. Wir haben gesehen, dass Husserl zufolge einige ernste Akte gänzlich ernst sind: wer sieht, phantasiert nicht.

---

44. Diese Versuche gehen meist zurück auf (Walton 1973, 1978).

45. Z. B. (Strawson 1974).

Alle Thesen zum Wesen des *Annehmens*, die wir hier diskutiert haben, können eigentlich erst bewertet und verstanden werden, wenn wir die Beziehungen des Annehmens zu den *Annahmen* in Betracht ziehen. Wie das Urteilen in einer intimen Beziehung zum deklarativen Satz, zum „Darstellungssatz“ (Bühler), manchmal leider auch „Urteil“ genannt, und so zu Propositionen, steht, so auch das Annehmen zum Annahmesatz und zur Annahme.<sup>46</sup> Ich möchte Peters Geduld nun aber nicht weiter strapazieren, da er sowieso nicht an Annahmen und Propositionen glaubt.<sup>47</sup>

#### LITERATUR

- Broad, C. D. 1913: „Critical Notice of *Über Annahmen*“. *Mind* 22, 90–102.
- Cairns, Dorion 1973: „Perceiving, Remembering, Image-Awareness, Feigning Awareness“. In: Fred Kersten & Richard Zaner (Hg.), *Phenomenology: Continuation and Criticism*. Den Haag: Martinus Nijhoff, 251–262.
- Cohen, L. Jonathan 1992: *An Essay on Belief and Acceptance*. Oxford: Clarendon Press.
- Dummett, Michael 1973: *Frege – Philosophy of Language*. London: Duckworth.
- 1993: „Mood, Force, and Convention“. In: ders., *The Seas of Language*. Oxford: Clarendon Press, 202–223.
- Eckermann, Johann Peter 1836–1848 : *Gespräche mit Goethe in den letzten Jahren seines Lebens*. Leipzig: Brockhaus.
- Edgington, Dorothy 1995: „On Conditionals“. *Mind* 104 (414), 235–329.
- Findlay, J. N. 1963<sup>2</sup>: *Meinong's Theory of Objects and Values*. Oxford: Clarendon Press.
- Gardies, Jean-Louis 1985: *Rational Grammar*. Munich: Philosophia Verlag.
- Gerrans, Philip & Mulligan, Kevin 2013: „Immaginazione, *default thinking* e incorporamento“ (ital. Übers. v. „Intentional Imagination and Delusion“). In: Daniela Tagliafico (Hg.), *Nuove teorie dell'immaginazione. Rivista di Estetica*, n.s., 53, 55–87.

---

46 Vgl. (Heller 1929). Zum Wesen der Beziehung zwischen Satzarten, geistigen Zuständen oder „Akten“ und illokutiver Kraft hat es eine Debatte zwischen Dummett und Davidson gegeben (Dummett 1993). Zur Vielfalt solcher Beziehungen, vgl. (Gardies 1985, 155–175).

47. Diese Arbeit wurde unterstützt durch das „Seventh Framework Programme FP7/2007-2013“ der Europäischen Kommission („grant agreement“ Nr. FP7-238128) und durch den schweizerischen NCCR für die affektiven Wissenschaften. Mein Dank geht an Reinhild Mulligan, Marion Hämmerli, Maik Sühr, Anne Desler, Simone Neuber, Anita Konzelmann, Jan Plate und Wolfgang Künne für ihre grosszügige Hilfe.

- Heller, Ernst 1929: „Zur Logik der Annahme (1. Artikel)“. *Jahrbuch für Philosophie und phänomenologische Forschung* 10, 485–513.
- Hertz, Paul 1931: „Vom Wesen des Logischen, insbesondere der Bedeutung des modus Barbara“. *Erkenntnis* 2, 369–392.
- 1935: „Über das Wesen der Logik und der logischen Urteilsformen“. *Abhandlungen der Fries'schen Schule*, Neue Folge, 6/2, 225–272.
- Husserl, Edmund 1950: *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, Erstes Buch*. Hua III. Hg. Walter Biemel. Den Haag: Martinus Nijhoff.
- 1956: „Persönliche Aufzeichnungen“. *Philosophy and Phenomenological Research* 16, 293–302.
- 1973: „Zur Phänomenologie der Intersubjektivität“. In: ders., *Texte aus dem Nachlass, Erster Teil: 1905-1920*. Hua XIII. Hg. Iso Kern. Dordrecht: Kluwer.
- 1979: *Aufsätze und Rezensionen (1890–1910)*. Hua XXII. Hg. Bernhard Rang. Den Haag: Martinus Nijhoff.
- 1980: „Phantasie, Bildbewusstsein, Erinnerung. Zur Phänomenologie der anschaulichen Vergegenwärtigungen“. In: ders., *Texte aus dem Nachlass (1898-1925)*. Hua XXIII. Hg. Eduard Marbach. Dordrecht: Kluwer.
- 1984: *Logische Untersuchungen*. Hua XIX/1. Hg. Ursula Panzer. Den Haag: Martinus Nijhoff.
- 1984a: *Logische Untersuchungen*. Hua XIX/2. Hg. Ursula Panzer. Den Haag: Martinus Nijhoff.
- 1993: *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*. Ergänzungsband. In: ders., *Texte aus dem Nachlass 1934-1937*. Hua XXIX. Hg. Reinhold N. Smid. Dordrecht: Kluwer.
- 2000: „Aktive Synthesen“. Aus der Vorlesung „Transzendente Logik“ 1920/21. Ergänzungsband zu „Analysen zur passiven Synthesis“. Hua XXXI. Hg. Roland Breeur. Dordrecht: Kluwer.
- 2002: *Logische Untersuchungen, Ergänzungsband Erster Teil*. Entwürfe zur Umarbeitung der VI. Untersuchung und zur Vorrede für die Neuauflage der Logischen Untersuchungen (Sommer 1913). Hua XX/1. Hg. Ullrich Melle. Dordrecht: Kluwer.
- Jaśkowski, Stanislaw 1967: „On the Rules of Suppositions in Formal Logic“. In: Storrs McCall (Hg.), *Polish Logic 1920–1939*. Oxford: Oxford University Press, 232-258. (Erstveröffentlichung 1934 in *Studia Logica* 1, 5–32.)
- Künne, Wolfgang 1995: „Some Varieties of Thinking. Reflections on Meinong and Fodor“. *Grazer Philosophische Studien* 50, 365–395.
- 1996: „Thought, Speech, and the ‘Language of Thought’“. In: Christian Stein & Mark Textor (Hg.), *Intentional Phenomena in Context; Papers from the 14th Hamburg Colloquium on Cognitive Science*, Bericht Nr. 55, Juni 1996, 53–90.

- Lemmon, Edward John 1965: *Beginning Logic*. London: Nelson.
- Mackie, John Leslie 1973: *Truth, Probability and Paradox: Studies in Philosophical Logic*. Oxford: Clarendon Press.
- Mally, Ernst 1912: „Gegenstandstheoretische Grundlagen der Logik und Logistik“. Ergänzungsheft zu Bd. 148 der *Zeitschrift für Philosophie und philosophische Kritik*.
- Martin, Michael G. F. 2002: „The Transparency of Experience“. *Mind & Language* 17, 376–425.
- Meinong, Alexius 1902: *Ueber Annahmen*. Leipzig: Verlag Johann Ambrosius Barth.
- 1965: *Philosophenbriefe. Aus der wissenschaftlichen Korrespondenz von Alexius Meinong*. Hg. Rudolf Kindiner, Graz: Akademische Druck- u. Verlagsanstalt.
- 1968: „Über emotionale Präsentation“, *Abhandlungen zur Werttheorie*. Gesamtausgabe Bd. 3, 283–467.
- 1973: „Ueber die Erfahrungsgrundlagen unseres Wissens“. In: ders., *Ueber philosophische Wissenschaft ... und andere Werke*. Gesamtausgabe Bd. 5, 367–481.
- 1977: *Über Annahmen*. Bearbeitet von Rudolf Haller, Gesamtausgabe Bd. 4, Graz: Akademische Druck- u. Verlagsanstalt.
- Mulligan, Kevin 1989: „Judgings: their Parts and Counterparts“. *Topoi, Supplement 2, La Scuola di Brentano*, 117–148.
- 1995: „Perception“. In: Barry Smith & David W. Smith, *The Cambridge Companion to Husserl*. Cambridge: Cambridge University Press, 168–238.
- 1999: „La varietà e l'unità dell'immaginazione“. *Rivista di Estetica, Percezione*, 53–67.
- 2009: „Was sind und was sollen die unechten Gefühle?“. In: Ursula Amrein (Hg.), *Das Authentische. Referenzen und Repräsentationen*. Zürich: Chronos Verlag, 225–242.
- 2010: „Husserls Herz“. In: Manfred Frank & Niels Weidtmann (Hg.), *Husserl und die Philosophie des Geistes*. Berlin: Suhrkamp, 209–238.
- 2011: „On Meaning Something and Meanings“. In: Benjamin S. Schnieder u. a. (Hg.), *Themes from Early Analytic Philosophy. Essays in Honour of Wolfgang Künne. Grazer Philosophische Studien* 82, 255–284.
- 2013: „Acceptance, Acknowledgment, Affirmation, Agreement, Assertion, Belief, Certainty, Conviction, Denial, Judgment, Refusal & Rejection“. In: Mark Textor (Hg.), *Judgement and Truth in Early Analytic Philosophy and Phenomenology* (History of Analytic Philosophy Series). London: Palgrave/Macmillan, 97–137.
- 2012: *Wittgenstein et la philosophie austro-allemande*. Paris: Vrin.
- Reinach, Adolf B. P. 1989: *Sämtliche Werke. Textkritische Ausgabe in 2 Bänden*. Hg. Karl Schuhmann & Barry Smith. München: Philosophia Verlag.

- Smith, Barry 1985: „Sur la modification du sentiment: l'esthétique de l'Ecole de Graz“. *Revue d'Esthétique* 9, 19–37.
- Stein, Edith 1970 (1922): „Psychische Kausalität“. In: dies., *Beiträge zur philosophischen Begründung der Psychologie und der Geisteswissenschaften. Eine Untersuchung über den Staat*. Tübingen: Max Niemeyer Verlag, 2–116.
- Strawson, Peter Frederick 1974: „Imagination and Perception“. In: ders., *Freedom and Resentment and Other Essays*. London: Methuen.
- Vendrell Ferran, Ingrid 2009: „Meinongs Philosophie der Gefühle und ihr Einfluss auf die Grazer Schule“. *Meinong Studien III/ Meinong Studies III*. Heusenstamm: Ontos Verlag, 199–239.
- 2010: „Ästhetische Erfahrung und Quasi-Gefühl“. In: Venanzio Raspa (Hg.), *The Aesthetics of the Graz School. Meinong Studies IV/Meinong Studien IV*, 129–168.
- Waismann, Friedrich 1976: *Logik, Sprache, Philosophie*. Stuttgart: Reclam.
- Walton, Kendall L. 1973: „Pictures and Make-Believe“. *The Philosophical Review* 82, 283–319
- Walton, Kendall L. 1978: „Fearing Fictions“. *The Journal of Philosophy* 75, 5–27.
- White, Alan R. 1990: *The Language of Imagination*. Oxford: Blackwell.
- Witasek, Stephan 1904: *Grundzüge der allgemeinen Aesthetik*. Leipzig: Johann Ambrosius Barth.
- 1923 (1907<sup>1</sup>): *Grundlinien der Psychologie*. Leipzig: Felix Meiner.
- Wittgenstein, Ludwig 1977: *Philosophische Untersuchungen*. Frankfurt: Suhrkamp.
- 1984: „Letzte Schriften über die Philosophie der Psychologie“. In: ders., *Bemerkungen über die Philosophie der Psychologie* (Werkausgabe Bd. 7). Frankfurt: Suhrkamp.



# MEANING, ENTERTAINING, AND PHANTASY JUDGEMENT<sup>1</sup>

Mark TEXTOR  
King's College, London

## *Summary*

Meinong takes assumption to be a *sui generis* attitude that plays an important role in our mental life that no other attitude or combination of attitudes can play. Why should one introduce such a *sui generis* attitude in the first place? Because, argues Meinong, it allows one to overcome a problem that arises for intentionalist theories of meaning that is otherwise not solvable. In this paper I will argue that Meinong's motivation for the introduction of assuming is convincing and, in turn, that one can shed light on assuming by seeing how it helps to develop a theory of meaning.

## 1. *Introduction*

Facts about linguistic meaning are not fundamental. The English sentence 'Germany will beat Brazil 7:1' means that Germany will beat Brazil 7:1, but it has this meaning in virtue of other facts. Which facts? According to the intentionalist view of meaning, an indicative sentence has the meaning it has in virtue of the fact that that it is regularly uttered with the intention that an audience comes to form or activate a particular belief (in virtue of recognizing this intention). However, frequently the utterer cannot expect and therefore not intend that her audience will form the required belief. For instance, if the utterer believes that the audience does not trust her, she cannot have the right intention, yet she can mean something by her

---

1. Central European Philosophers such as Bolzano, Brentano, and Twardowski have for a long time been unjustly neglected or even ignored. Peter Simon's work has helped to change this situation. He has made us aware of the importance of the Austrian and Polish contribution to the philosophy of logic and metaphysics and he has mined the works of these philosophers for new insights into old problems. Among other things, he has shown that Meinong's philosophy contains much more than things that don't exist. (See, for instance, Simons 1996, 1999, 2013.) I am therefore very happy to contribute this piece on Meinong to a volume in his honour.

utterance. If one wants to hold on to the intentionalist view, one needs therefore to identify a different intended response. Meinong argues that the intended response is that the audience *assumes* that p. (Note: Meinong's assumptions are not assumptions in the usual sense of the word. More in due course.) When she utters 'Germany will beat Brazil 7:1' the speaker intends that her audience makes the assumption that Germany will beat Brazil 7:1. If Meinong is right, Brentanians like Marty are wrong: not every mental phenomenon is either a presentation, i.e. an act of thinking of something, a judgement (belief) or an act of love or hate or a combination of these. In this paper I will develop and defend Meinong's response and in doing so unfold his concept of assumption.

The plan of this paper is as follows. In section two I will expound the problem for intentionalist views of meaning; in section three I will criticise Marty's and Grice's proposals to solve it. This will pave the way for Meinong's theory of assumptions (section 4) on which his solution of the problem is based (section 5). I will argue that Meinong's assumptions are phantasy judgements: one assumes that p by putting oneself inside the skin of someone who believes that p.

## 2. *Intentionalist semantics*

The Austrian philosopher Anton Marty (1847–1914) defended a version of the intentionalist view in his book *Untersuchungen zur allgemeinen Grammatik und Sprachphilosophie*.<sup>2</sup> He expounds the main claim as follows:

That a linguistic means has the meaning (or meaning function), say, of a statement [*Aussage*], means therefore: as a rule it is aimed (and, to some extent, also able) to suggest or trigger a certain judgement in the hearer. (Marty 1908, 286. Quoted and translated in Cesalli 2013, 148)

Later Grice (1957) will make this idea the basic building block of his theory of meaning and communication. He makes a good case that a speaker means that p by something she does only if she intends that the audience arrives at the belief that p by, in part, recognizing this very intention.

One can sum up the core idea of meaning intentionalism as follows:

---

2. See Cesalli (2013) for a detailed analysis of Marty's theory of meaning. See also Mulligan (2011).

(S-Meaning) A linguistic sign  $x$  is an indicative sentence that means that  $p$  if, and only if, it is a rule that utterances of  $x$  are made with the intention to invoke in an audience the judgement that  $p$ .

(S-Meaning) is a metaphysical thesis about which facts determine (i) whether something  $x$  is an indicative sentence and (ii) which meaning  $x$  has.

Both meaning and sentential character are grounded in a regularity that concerns an audience responding to utterances of  $x$  in a way intended by the speaker. Marty proposes that the intended response is that the audience makes a judgement. Grice (1989, 351) is adamant that the intended response to an indicative utterance is that the audience forms a belief. It is worth quoting him on this point:

The most general complaint, which comes from Strawson, Searle, and Mrs. Jack, seem to be that I have, wholly or partially misidentified the intended [...] effect in communication; according to me it is some form of acceptance (for example, belief and desire), whereas it should be held to be understanding, comprehension, or (to use an Austinian designation) “uptake”. ... [The objector’s position] hardly seems satisfactory when we see that it involves attributing to speakers an intention which is specified in terms of the very notion of meaning which is being analysed (or in terms of a dangerously close relative of that notion). Circularity seems blatantly abroad. (Grice 1989, 351)

If the meaning of signs is determined by the intentions with which they are regularly used by speakers in communication, what the speakers intend (the intended effect or response) cannot be that the audience understands or comprehends the uttered sign or that the audience achieves uptake of what the speaker said on the occasion of utterance. Otherwise one goes in a rather tight circle. For one understands the uttered sign if, and only if, one knows what it means. Similar points apply to such notions as knowledge of truth-conditions if the truth-conditions for a sentence  $s$  are supposed to reveal the meaning of  $s$ . Marty and Grice specify the intended response in terms of the acquisition of a propositional attitude and thereby avoid circularity.<sup>3</sup> In the next section we will see that Marty and Grice face another problem that is grist to the mill of their opponents.

---

3. Marty and Grice both (plausibly) assume that the content of propositional attitudes is determined prior to communication. However, they don’t provide an account which underwrites this assumption.

### 3. *The problem of the missing intended effect*

A problem with the view that the intended effect of an indicative utterance is that the audience makes a judgement is that the audience's response to an utterance must be within the audience's control (See Grice 1989, 221). If we are concerned with a speaker who means that  $p$  when making an indicative utterance, the intended effect cannot be merely caused. The audience must be able to decide or to refuse to  $\phi$ , where ' $\phi$ ' designates the intended effect. But a speaker has no (immediate) control over his judgements: people cannot literally decide or refuse to believe. You might utter 'Germany will beat Brazil 7:1', expressing your belief that Germany will beat Brazil 7:1, intending to bring about that I come to judge that this is so by recognizing this very intention, but I simply cannot bring myself to believe that Germany will beat Brazil 7:1. I take this possibility to be beyond belief. Hence, the intended effect will often not come about. Speakers know that this is so and can therefore, in general, not (rationally) intend to bring about the effect. For one can only rationally intend to achieve an effect by doing  $X$  if one believes that there is a chance that doing  $X$  will produce the effect.

Marty is aware of these problems for his form of meaning intentionalism. The audience will only make the intended judgement if it takes the speaker to be truthful and to know what he communicates. (Marty 1908, 206.) But a speaker can certainly mean that  $p$  by an utterance even if these conditions are not met. Marty's response to our problem is that the speaker's intention in uttering a linguistic means need not be successful:

We said that the primary intention of the speaker is to produce an analogous judgement (with respect to content and quality) in the hearer as that which is expressed by the statement as a rule. *But of course the success of this intention does not belong to an understanding of the statement.* Rather it suffices that the hearer arrives at a presentation of the content of that judgement which the statement is (usually) intended to invoke. (Marty 1908, 362; my translation, my emphasis)

We should not read Marty as using here the notion of understanding. This would undermine the very point of meaning intentionalism, Instead we can and should read him as proposing that the intended effect of an utterance need not be achieved for the speaker to mean something by it. It suffices that the audience forms a presentation of the content of the belief the speaker intends to produce in virtue of recognizing the speaker's intention that they should adopt this belief. For example, faced with an

utterance intended to invoke the judgement that *p*, the audience might find that *p* 'beyond belief'. Nonetheless they can still think: 'That *p* is what the speaker wants me to believe. But, well, I simply can't believe that *p*.' In this situation, Marty argues, the speaker has meant that *p* by his utterance, although the intended effect of his utterance was not achieved.

Marty's response to the problem of the missing intended effect is unsuccessful for two reasons.

*First*, there are many presentations of the propositional content that Germany will beat Brazil 7:1. Which one does a speaker need to form to count as understanding the speaker's utterance? We need an answer to this question since not every presentation of the propositional content enables one to understand the utterance. For example, thinking of the propositional content merely as *the content the speaker wants to express* will not suffice. This problem would not even arise if the audience could have an attitude with the content that Germany will beat Brazil 7:1. But Marty's psychology has no room for such an attitude. Like other Brentanians, he takes every mental phenomenon either to be a presentation, a negative or positive judgement or an act of love or hate. The response of the audience is not supposed to be a judgement and it is implausible that it is an act of love or hate. Hence, according to Marty, it can only be a presentation of the propositional content. But he provides no independent principle that singles out how the speaker is supposed to present this content.

*Second*, it is not the success of the intention, but the formation of the intention that is the basic problem. The speaker can only intend to produce the response in his audience by something he does if he believes that there is at least a chance that his doing so will produce the response. In many situations the speaker does not expect to be able to induce the required belief in his audience. More importantly, often the speaker believes that she will *not* be able to do so. If you are notorious for false alarms and know that you are notorious for false alarms, you cannot expect that your audience will come to form the belief that your flock is being attacked by a wolf when you utter 'Wolf!'. Hence, you cannot form the intention necessary to mean that a wolf is attacking your flock by an utterance, namely the intention to induce the required belief in your audience. Marty's proposal ignores the speaker side of the problem and therefore fails.

Grice proposed a different solution to the problem of the missing effect that is free of the problems just mentioned.<sup>4</sup> There are cases where it is

---

4. See Grice (1989, 111; 123) on protreptic utterances.

independently plausible that the speaker intends that the audience forms a belief *about the speaker's beliefs*. Consider an example from Grice: If a student answers the examination question whether Napoleon won the battle of Austerlitz by uttering 'Napoleon won the battle of Austerlitz' with assertoric force, the student's utterance is made with the intention that her examiners judge that the student knows that Napoleon won the battle of Austerlitz. Normally, the student will neither intend that the examiners acquire nor that they activate their belief that Napoleon won the battle of Austerlitz. Similarly, when you utter 'Germany will beat Brazil 7:1', you may not be able to intend that your audience comes to believe that Germany will beat Brazil 7:1 by recognizing your intention, but you can intend that the audience comes to believe that you, the utterer, believe that Germany will beat Brazil 7:1. The speaker intends, argues Grice, that the audience arrives at the belief that the speaker herself believes that p. In favourable circumstances, taking the speaker to believe that p will give the audience a reason to come to believe that p.

Marty criticized and rejected this response roughly sixty years before Grice developed it in print. For Marty continues the quote on p. 288 as follows:

Also the belief that someone now actually intends to insinuate such a judgement, further, the assumption that the speaker—if he is known—indeed has the belief which is usually expressed by the statement is not essential. I can understand the statement although I explicitly recognize that it is a thoughtless or untruthful (lying) utterance and I talk about understanding a sentence even when I don't know that the utterance under consideration is actually made by someone. (Marty 1908, 362; my translation, my emphasis)

Marty sets Grice's solution aside because the formation of a belief about the belief the speaker holds is not essential for 'understanding' in the sense of successful communication. Meinong agrees with Marty on this point:<sup>5</sup>

Sometimes it indeed happens that, when X expresses a belief addressing Y, the latter judges: "X believes that ...",—or more precisely: "There is a judgement of the subject X with an object that is so and so." But to think this thought really through, especially to bring clearly to mind the mental state of affairs that holds of X with respect to its act and object and make a judgement about it, is by far too difficult a task as that one can expect a psychologically untrained person to master it, in particular as an easy everyday task. (Meinong

---

5. See also McDowell (1980, 38ff).

1910, 39 [34]. My translation (all Meinong translations are mine). References to the pagination of the English translation in square brackets.)

Meinong's reason for arguing that the audience's response to an indicative utterance cannot in general be a judgement about the speaker's belief is, however, nothing more than an appeal to intuition. Is it not rather easy for most of us to come to believe that the speaker believes that *p*? But the principled problem for the view that the intended response is a judgement has not been finessed by making the intended judgement one about the speaker's belief. For it is not always open to one to form a belief about the belief of a speaker. If the hearer openly mistrusts the speaker, the speaker cannot expect and therefore cannot intend that the hearer acquires the belief that the speaker himself believes what he says.<sup>6</sup>

#### 4. *Meinong on assuming*

If you can learn a simple trick, Scout, you'll get along a lot better with all kinds of folks. You never really understand a person until you consider things from his point of view, until you climb inside of his skin and walk around in it.

Atticus Finch

The negative points made so far give us a blueprint for what the intended effect of an indicative utterance must be like. The intended response should:

1. be a propositional attitude (and not a presentation of a propositional content);
2. not involve the notion of meaning in its content;
3. be under the control of the audience and independent of what it believes;
4. not be an attitude about an attitude.

Is there an attitude that fits this description? Neale (1992, 547) proposes that actively entertaining that *p* is the intended effect. If this proposal is on the right track, the notion of entertaining is crucial for articulating the intentionalist theory of meaning. However, entertaining seems, so far, to be merely be a label for an attitude that fits our description. Kriegel

---

6. See Rumfitt (1995, 834).

(2013, 16), in my view correctly, remarks that ‘our present understanding of entertaining is surprisingly patchy’.

Meinong aims to complete our understanding of what is now often called ‘entertaining’. He (1910, 6 [12, Fn. 6]) credits Frege’s (1891) with discovering what he himself calls ‘assumptions’. But while Frege is content with acknowledging that one must be able to merely think or entertain a proposition without commitment to its truth, Meinong is to my knowledge the first philosopher who recognised the central role of entertaining in the theory of meaning *and* provided an account of it. He writes:

[Assumptions] have a dominating, often, so to speak, already outwardly recognizable function in understanding speech and writing (as the bearer of sentence in contrast to mere word meaning), games and art, and in motivational desire [...]. (Meinong 1921, 31 [89])

In the following sections I will outline the main points of Meinong’s account of assumption and go on to argue that assumption is the propositional attitude that will help us get the intentional view of meaning off the ground.

So what is an assumption or, more fundamentally, what is it to assume something? First, a note of caution. The word ‘assumption’ is misleading.<sup>7</sup> Ryle makes this point nicely:

[I] want to suggest, what is certainly debatable, that the doctrine of Meinong is anyhow partly right, that there is a sort of “thinking that” which is completely neutral and attitude-less, and that this must be present as an ingredient in all the varieties of “thinking that.” Meinong’s word “Annahmen” has been translated “assumption” and “supposal,” which both suffer from not being neutral enough. As regards the verb, I prefer to speak of “entertaining” or “imagining that” or, best of all, “thinking of ... as.” This seems to me the common element in all the other thinking attitudes. (Ryle 1929, 113)

Meinong’s assumptions are neither hypotheses nor suppositions. Ryle’s ‘entertaining’ and especially ‘imagining’ point us in the right direction. However, Ryle’s ‘thinking of ... as’ is less helpful. When I entertain the thought that it often rains, I don’t think of something as something. At the end of this section we will see that Meinong has a terminological alternative to ‘assumption’ that is free of the problems Ryle detected. We will also see that Ryle is wrong when he says that Meinongian assuming is a common ingredient in all varieties of ‘thinking that’.

---

7. See also Künne (1996, 368).

In *On Assumptions*, Meinong motivates the introduction of assumption as a new propositional attitude by first looking at games and then moving on to art. Let us retrace Meinong's first steps in order to answer our own questions.

The natural home of assuming is the practice of playing games of make-believe. When children play cowboys and indians, they *act as if they believe* that, for example, a chair is a horse, in the full knowledge that the chair is not a horse. They don't believe that the chair is a horse; neither do they deceive themselves when playing, and nor do they merely think of a situation in which a chair is horse. So what do they do? Meinong's first stab answer seems merely to put a label on the attitude in question:<sup>8</sup>

The intellectual behavior of the player is less than judgement, but it is more than presenting because it is assuming. (Meinong 1910, 112 [84])

How can we get a better grip on assuming?

Sometimes Meinong uses the 'less than judgement' characterization to shed light on assuming. He takes every judgement to have the quality of belief: someone who judges that *p* is convinced that *p*. He tentatively suggests that an assumption is a judgement without the quality of belief or a limiting case of judgement: it is a judgement with zero degree of belief. (Meinong 1910, 340 [242] and 344 [245].) However, I can assume that  $2 = 1$ , even though I know that  $2 = 1$  is impossible. To describe assuming that  $2 = 1$  as judging that  $2 = 1$  with a zero degree of belief is introducing a Pickwickian sense of 'to judge'.<sup>9</sup>

How can one, then, say more about assuming?<sup>10</sup> Meinong uses the locution 'sich einbilden [...] zu sein' to talk about the act of assuming. I think the best translation of this phrase is 'to fancy', with a simple complement, or 'to be'. Meinong's model for *fancying oneself to be* is acting:

If the actor is able to "fancy himself" [sich einbilden zu vermag] to be the person he has to play and himself to be in the environment given by the plot of the play, he will, given sufficient talent on his side, act outwardly in a way the person played is likely act [...]. (Meinong 1910, 114 [85])

The actor playing Nelson Mandela fancies himself to be Nelson Mandela and to be in the circumstances described in the play. The children play-

---

8. See also See Russell (1904, 340).

9. See also Künne (1996, 371).

10. Walton (1990), 20 talks about imagination and imagining oneself to be, where Meinong talks about assuming. Walton leaves imagination unanalyzed.

ing cowboys and indians fancy themselves to be cowboys and to be in a world that is as the players have agreed it to be, where, for example, the chair is a horse.<sup>11</sup>

*Fancying oneself to be* is further glossed as *putting oneself in the situation of someone else*. We have all from an early age developed the ability to put ourselves in the position of someone else (Meinong 1910, 115 [85]). Actors cultivate the ability to put themselves into the shoes of another person, but, as Atticus Finch suggests, we all have this ability and should exercise it. This ability holds the key to understanding what an assumption is:

This “to put oneself in the position” is known as the fact of assumption [...]. (1910, 119 [89])

The connection between assuming and putting oneself in the position of someone deserves to be highlighted. Neither Russell in his review of the first edition of *On Assumptions*, nor Broad in his critical notice of the second edition picked up on the fact that Meinong takes assuming to be an upshot of the ability to put oneself in the position of someone else. Kroon (1992, 505ff.) recognises the importance of the ‘putting oneself in the position’ formula. But he focuses on cases of pretence, yet not all cases of putting oneself in the position of someone else are cases of pretence. I can put myself in the position of my neighbour in order to understand why he is upset about the wall I am building, but in doing so I am not pretending to be my neighbour.<sup>12</sup>

Meinong explains assuming by appeal to the ‘putting oneself in the position of’ formula:

*T* assumes that *p* if, and only if, *T* puts himself in the position of someone who believes that *p*.

Consider the experienced liar. The liar who wants to deceive his audience needs to present himself as someone who himself believes that *p*. How can

---

11. According to Twardowski, the actor on stage is supposed to present, that is, think of feelings; the product of this activity is then an artifact that looks like the behavior of a person who has these feelings. But if a presentation of a feeling is a mental act that is directed on the feeling this is highly implausible. Lee Strasberg would not be happy if you were merely thinking of longing for freedom when you tried to play Nelson Mandela. Twardowski comments that, ‘[i]f one prefers here to talk instead of presentations of emotions [Gefühlsvorstellungen] of phantasy emotions [Phantasiegefühlen], this will not change anything essential’ (Twardowski 1912, 59). In fact the change seems essential and to be a step in the right direction.

12. See Sainsbury (2010, 10f.).

he do this if he does not believe that p? The answer is that he puts himself in the position of someone who believes that p. This allows him to act *convincingly* as if he believed that p and thereby take others in. (Meinong 1910, 119 [89].)

Assuming that p is conceptually dependent on judging (believing). This dependence comes out in an analogy Meinong uses to explain the nature of assuming:

Assumptions are related to judgements as phantasy presentations are related to serious presentations and can therefore also be called *phantasy judgements*. (Meinong 1921, 31 [89])

Meinong takes the term ‘phantasy judgement’ to shed light on the nature of assumptions. He coins this term in analogy with his term ‘phantasy feeling’:

If one transfers this designation [*Phantasiegefühl*] to the realm of thinking, one arrives at the expression “phantasy judgement” for “assumption”, and, as far as I can see, this is an expression which is well understood by someone who hears it for the first time and it conveys something determinate that characterizes the assumption facts correctly [...]. (Meinong 1910, 383 [273])

In order to understand the ‘phantasy judgement’, we need first to get a grip on ‘phantasy feeling’. So what is a phantasy feeling?

*Prima facie*, one is scared when one sees a horror movie. But if one were really scared, why doesn’t one leave the cinema? Presumably because these feelings are not real, but are mere phantasy emotions. ‘Phantasy’ modifies ‘emotion’: a phantasy emotion is not an emotion, because it lacks motivational force. (Meinong 1902, 236.) But a phantasy emotion is sufficiently similar to an emotion such that everyone who understands the term ‘emotion’ understands what it means to say that one feels horror when one watches the movie. (ibid.) Similarly, everyone who understands the verb ‘see’ understands an utterance of ‘I saw pink rats’, although one can be fully aware that one cannot see what is not there. (See Grice 1989, 44.) Just as one can use ‘see’ in a relaxed sense to describe certain visual hallucinations, one can use ‘feel’ and ‘emotion’ in a relaxed sense to describe what one experiences when one watches a horror movie. A phantasy emotion is sufficiently like an emotion from the inside to warrant the title ‘emotion’ or ‘feeling’, yet it is unlike an emotion in terms of its motivational effects.<sup>13</sup> A phantasy judgement or assumption is like a phantasy emotion;

---

13. Walton (1990, 202) agrees with Meinong. Sainsbury (2009, 18) argues that a phantasy emotion is not only like an emotion; it is a pathological emotion (like arachnophobia) gener-

it is a mental event that is sufficiently like a judgement from the inside, yet lacks the motivational force of judgements. Why is assuming like judging from the inside? If one puts oneself in the shoes of someone who judges that *p*, one does not oneself judge that *p*, but one will attend to considerations that support that *p*, etc. Meinong's terminological recommendation solves the problems connected with 'assumption'. I will therefore follow him and use the term 'phantasy judgement'.

Let us take stock. Meinong gives us a good reason to hold that there is a basic psychological ability that we possess from an early age onwards: we are able to put ourselves in the position of someone else. This ability is exercised for different reasons when we are playing games and acting. Thus a complete philosophical psychology needs to recognize phantasy judgement, in addition to judging, presenting, and feeling. Phantasy judging is a propositional attitude that is under our control: we can decide to put ourselves in the position of someone who believes something or we can refuse to do so. We can do so independently of what we believe or believe to be possible. This seems, therefore, to be the attitude we are looking for. However, phantasy judging is not, as Ryle proposes, the common ingredient in all propositional attitudes. The judgement that *p* is not a phantasy judgement plus a further property—say, a degree of belief that exceeds 0. Nor is a phantasy judgement a judgement minus the belief feature. Just as the notion of a fake banknote depends on the notion of a banknote, the notion of phantasy judgement conceptually depends on the notion of judgement. And just as a fake banknote is not a kind of banknote, a phantasy judgement is not a kind of judgement.<sup>14</sup>

### 5. *Meaning, intended effect and phantasy judgement*

Equipped with the notion of phantasy judgement, Meinong returns at the end of *On Assumptions* to the problem we set out in section 2. He writes that:

Insofar as one aims to state precisely the minimal condition of when one can say that a discourse has been understood one can simply state: the hearer

---

ated by vivid representations. Since phantasy emotions are not my main topic here I will not pursue this matter further.

14. The idea that imaginings are voluntary acts which in important respects are like perceptions, judgements and desires is explored in Currie and Ravenscroft (2002, 11).

understands when he assumes what the speaker says. Just as the achievement of the spoken sentence consists at least in expressing an assumption, the achievement of the heard sentence consists at least in evoking an assumption in the hearer. [...] However, it is always more distinct and precise to say the following: understanding of something spoken (naturally also of something written) consist in grasping its meaning. (Meinong 1910, 361 [258])

I read Meinong as addressing two questions without clearly distinguishing them. The problem we identified in section 2 was the problem how the intended effect of an indicative utterance can be identified without making use of the notion of meaning. He speaks to this problem, but he also addresses the question under which conditions one understands a sentence when it has already been invested with meaning. Let us distinguish the questions that Meinong seems to run together, and consider first his answer to the second question.

What is it to understand an indicative sentence of a language that means that p? Meinong, in essence, answers that:

(S-Understanding) If the indicative sentence *s* says that p, one understands *s* if, and only if, one grasps that p.

What is it to *grasp* the meaning of a sentence? In the paragraph before the quotation above Meinong explains that:

Understanding a sentence simply consists in grasping an objective and it does not matter whether this grasping is made by way of judging or assuming. (ibid.)

An objective is, roughly, a state of affairs. So if the meaning of the sentence 'Germany will beat Brazil 7:1' is the state of affairs that Germany will beat Brazil 7:1, one grasps the meaning of this sentence if, and only if, one is either disposed to judge or to phantasy judge that Germany will beat Brazil 7:1.

This characterization of sentence understanding is unsatisfactory. *Prima facie*, neither being disposed to phantasy judge nor being disposed to judge that p suffices for understanding a sentence that says that p. For I must be so disposed for the right reason. If it is a brute regularity or association that I come to judge or phantasy judge that p when hearing a particular sentence, I don't thereby manifest my understanding. In part, one should

be disposed to make these judgements or phantasy judgements because one knows what the sentence means, that is, if I know the true answer to the question ‘What does *s* mean?’<sup>15</sup> But then the notion of knowledge of meaning does all the interesting work.

I think therefore that Meinong’s notion of phantasy judgement does not help to articulate our notion of understanding a sentence. Its value lies in the fact that it enables us to specify the intended effect of an indicative utterance in a satisfactory way. In short:

(S-Meaning<sub>Meinong</sub>) A linguistic sign *s* is an indicative sentence that means that *p* if, and only if, it is a rule that utterances of *s* are made with the intention to invoke in an audience the *phantasy judgement* that *p* (by the audience’s recognition of this intention).

The basic idea of (S-Meaning<sub>Meinong</sub>) is that the speaker openly intends the audience to, as one might intuitively put it, put themselves in the position of someone who believes that *p*. The audience might then, in a further step, come to acquire the belief that *p*. Is this a promising proposal? Yes, for the following four reasons:

*First*, phantasy judging is under the control of the speaker. I can decide to phantasy judge or refuse to do so in the light of your utterance. Hence, (S-Meaning<sub>Meinong</sub>) does not require the meaning intentionalist to implausibly conceive of the intended effect as a second-order attitude. Meinong solves the basic problem of the intentional view by identifying the intended effect of an indicative utterance with a phantasy judgement. Grice (1989, 384) himself gives an example in which the intended effect is not under the audience’s control. I can do something intending to make you distressed by recognising my intention to distress you. But it is not under your control to be distressed or not. Grice has no natural treatment of this problem; Meinong has one ready to hand. The intended effect is a phantasy feeling.

*Second*, while the immediate intended effect of an indicative utterance is a phantasy judgement, the mediate effect is often belief formation or activation.<sup>16</sup> How are the immediately intended effect and mediately intended effect related? Meinong’s view of phantasy judgement allows him to answer this question. A hearer who puts herself in the position of

---

15. For a dissenter see Pettit (2002).

16. See Grice (1989, 352) and Neale (1992, 547).

someone who believes that *p* will acquire the belief herself easily if nothing counts against it. For example, Meinong observes that liars often tend to believe their own lies (Meinong 1910, 119 [89]). The liar has put himself in the position of someone who believes that *p*, and ends up actually believing that *p*. There is therefore a smooth transition from phantasy judgement to judgement; that is, in the case of an indicative utterance, from immediately intended to mediately intended effect. This smooth transition helps to explain why communication is beneficial: in the good case we acquire beliefs that constitute knowledge, without requiring that the immediate intended effect is a judgement.

*Third*, the putting-oneself-in-a-position mechanism helps to explain the appeal of the intentionalist view. In communication, the suggestion goes, we exploit the putting-oneself-in-a-position mechanism. The speaker reveals that he believes that *p* with the intention that the audience put themselves in the position of someone who believes that *p*.

*Fourth*, although one's understanding a sentence that means that *p* does not consist in being disposed to judge or phantasy judge that *p*, the putting-oneself-in-a-position mechanism is still in play in communication involving such sentences. Consider an example from Stalnaker (2002, 718-9). *A* utters 'The man drinking a martini owns an Aston Martin'. You know that the man *A* is talking about drinks water from a martini glass. If you want to keep the conversation going without a distracting correction ('You mean, I take it, the man drinking from a martini *glass*') you will simply *assume* that the man with the martini glass owns an Aston Martin. The conversation proceeds then on the basis of a shared *assumption*, not on the basis of a shared belief.

Stalnaker himself describes this example as involving acceptance of a proposition:

To accept a proposition as true is to treat it as true for some reason. (Stalnaker 2002, 716)

Meinong's assuming is more fundamental than acceptance. One treats a proposition as true if, and only if, one assumes it, that is, if *one puts oneself into the position of someone who believes that p*. We can do this for various reasons. The case Stalnaker describes is one in which one assumes a proposition because one wants to continue the conversation.

Let us finally address a basic worry. Phantasy judging is, as we have seen, conceptually dependent on judging. If one needs to explain judgement by appealing to the concept of truth, the intentionalist view is undermined.

If it needs to rely on the concept of truth, why not let truth and truth-conditions do the work in the theory of meaning?

Meinong himself does not help us to answer the threat of circularity. In *On Assumptions* he takes judgement to be primitive. But it is open to him to argue that our grip on belief and desire is due to their role in the explanation of behaviour. It is therefore at least not obvious that we need to appeal to truth to make sense of judgement and assumption.

To conclude: Equipped with Meinong's notion of assumption, one can develop a version of intentionalist semantics that overcomes a serious problem for this view. Whether further problems can be solved in similar fashion is a question I shall leave for another occasion.<sup>17</sup>

#### REFERENCES

- Brentano, Franz 1874: *Psychologie vom Empirischen Standpunkt*. (Two volumes). Hamburg: Felix Meiner 1955. Trans. by Linda MacAlister: *Psychology from An Empirical Standpoint*. London: Routledge 1995.
- Broad, Charles D. 1913: "Critical Notice of A. Meinong *Über Annahmen*, 2<sup>nd</sup> edition". *Mind* 22(85), 90–102.
- Cesalli, Laurent 2013: "Anton Marty's Intentionalist Theory of Meaning". In: Guillaume Frechette & Denis Fisette (eds.), *Themes from Brentano*. Amsterdam: Rodopi, 139–63.
- Currie, Gregory & Ravenscroft, Ian 2002: *Recreative Minds*. Oxford: Oxford University Press.
- Frege, Gottlob 1891: "Funktion und Begriff." Jena: Herman Pohle.  
Translated as "Function and Concept". In: Brian McGuinness (ed.), *Gottlob Frege: Collected Papers on Mathematics, Logic and Philosophy*. Oxford: Basil Blackwell, 137–56.
- Grice, Paul 1957: "Meaning". *The Philosophical Review* 66(3), 377–88.  
— 1989. *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.
- Kriegel, Uriah 2013: "Entertaining as a Propositional Attitude: A Non-reductive Characterisation". *American Philosophical Quarterly* 50(1), 1–22.

17. I presented a previous version of this material at the workshop *Théorie des Actions et des Produits de Kazimierz Twardowski in Paris in 2014*. I am grateful to Jacques Dubucs, Giulia Felappi, Bob Hale, Nick Jones, Uriah Kriegel, Nils Kürbis, Sandra Lapointe, Jessica Leech, Guy Longworth, Jerold Levinson, and Göran Sundholm for comments that led to significant changes.

- Kroon, Frederick 1992: "Was Meinong only Pretending?". *Philosophy and Phenomenological Research* 52(3), 499–527.
- Künne, Wolfgang 1995: "Some Varieties of Thinking: Reflections of Meinong and Fodor". *Meinong and the Theory of Objects. Grazer Philosophische Studien* 50, 365–395.
- Longworth, Guy 2008: "Linguistic Understanding and Knowledge". *Nous* 42(1), 50–79.
- Marty, Anton 1908: *Untersuchungen zur Allgemeinen Grammatik und Sprachphilosophie* Bd.1. Halle: Niemeyer.
- Mulligan, Kevin 2011: "On Meaning Something and Meanings". In: *Themes from Early Analytic Philosophy. Grazer Philosophische Studien* 82, 255–284.
- McDowell, John 1980: "Meaning, Communication and Knowledge". In: Zak van Straaten (ed.), *Philosophical Subjects: Essays on the Work of P. F. Strawson*. Oxford: Clarendon Press, 117–39. (Reprinted in his: *Meaning, Knowledge and Reality*. Cambridge, Mass.: Harvard University Press 2001.)
- Meinong, A. 1902: *Über Annahmen*. Leipzig: Johan Ambrosius Barth.
- 1910: *Über Annahmen*, 2<sup>nd</sup> edition. In: *Alexius Meinong Gesamtausgabe IV*. Graz: Akademische Druck und Verlagsanstalt 1977. (Translated by James Heanue as: *On Assumptions*. Berkeley: University of California Press 1983.)
- 1921: *Selbstdarstellung*. In: *Die Deutsche Philosophie der Gegenwart in Selbstdarstellungen*. Bd. I. Leipzig: Felix Meiner. (Reprint in his: *Über Gegenstandstheorie/ Selbstdarstellung*. Hamburg: Felix Meiner 1988.)
- Neale, Stephen 1992: "Paul Grice and the Philosophy of Language". *Mind and Language* 15(5), 509–559.
- Pettit, Dean 2002: "Why Knowledge is Unnecessary for Understanding Language". *Mind* 111 (443), 519–50.
- Rumfitt, Ian 1995: "Truth-Conditions and Communication". *Mind* 104(416), 827–62.
- 2006: "Meaning and Understanding". In: Frank Jackson & Michael Smith (eds.), *Oxford Handbook of Contemporary Philosophy*. Oxford: Oxford University Press, 427–453.
- Russell, Bertrand 1904: "Review of Meinong *Über Annahmen*". *Mind* 13(51), 336–54.
- Ryle, Gilbert 1929: "Are there Propositions?". *Proceedings of the Aristotelian Society* 30, 91–126.
- Sainsbury, Mark 2010: *Fiction and Fictionalism*. London: Routledge.
- Simons, Peter 1996: "Meinong's Theory of Sense and Reference". *Meinong and the Theory of Objects. Grazer Philosophische Studien* 50, 171–86.
- 1999: "Bolzano, Brentano and Meinong. Three Austrian Realists". *Royal Institute of Philosophy Supplement* 44, 109–136.

- Simons, Peter 2013: "And Now for Something Completely Different: Meinong's Approach to Modality". *Humana Mente Journal of Philosophical Studies* 25, 119–134.
- Stalnaker, Robert 2002: "Common Ground". *Linguistics and Philosophy* 25(5-6), 701–25.
- Walton, Kendall L. 1990: *Mimesis as Make Believe: On the Foundations of Representational Art*. Cambridge, Mass.: Harvard University Press.

## THE THINGS WE CALL TRUE

Maria VAN DER SCHAAR  
Leiden University

### *Summary*

Peter Simons (2003) defends the thesis that both the act of judgement and the act of mere understanding are proper bearers of truth and falsity. There are several problems with this position, for example, a ‘true act of understanding’ cannot mean an *act of understanding, whose content is true*, for Simons does not acknowledge the notion of truth of propositional contents; therefore, it has to mean a *proper act of understanding*, but this is not the intended meaning.

It is argued that Twardowski’s distinction between the act of judgement and the judgement made is in need of a third notion, the judgement candidate as a bearer of truth and falsity. By means of this notion it is possible to answer the criticism raised against Simons’ position without invoking a Platonic notion of proposition.

### 1. *Introduction*

What are the things we call true? We speak of true sentences, true thoughts, true beliefs, the true religion, or of Peter Simons as a true friend, and a true inspiration for me since 1989. There are thus many kinds of things we call true, so we need to ask a more specific question. Assuming that some things are called true in a derived sense, and other things are called true in a proper, original sense, we may ask: What are the things we properly call true?

Peter Simons proposes to take the act of judgement to be a proper bearer of truth and falsity. If one defends the thesis that the act of judgement is a proper bearer of truth, one needs to give an answer to the following questions:

1. How are we to explain the difference between asserted and unasserted propositions? If we call judgements proper bearers of truth,

- how can we explain that what is not judged, such as the mere thought that Dublin is in Ireland, can also be called true?
2. How can we explain the validity of *Modus Ponens*? How are we to explain the identity of  $p$  in  $\vdash p \rightarrow q$  and  $\vdash p$ ?
  3. Is the proper bearer of truth also a proper bearer of falsity?
  4. If it is defended that the proper bearer of truth and falsity is in space and time, how can one explain that truth is not relative to space and time? If the proper bearer of truth and falsity is an act of our mind, how can we explain that truth is not subjective?
  5. If the bearer of truth is a mental act, how can we speak of truths that are not actually thought of? And can we speak of truths that have never been thought of?
  6. What is the relation between judgments and assertings as bearers of truth and falsity and other speech acts such as questions and orders?
  7. What does the *that*-clause refer to in ‘That snow is white is true’?

The term ‘judgement’ is ambiguous. It may denote the act of judgement, or its product, the judgement made, but it may have other meanings besides. As I will argue in section 4, we are in need of a third notion, the judgement candidate, the judgement that may not yet be judged, but is apprehended nonetheless. The judgement candidate can either be understood as a declarative sentence type together with its meaning, or as the meaning in abstraction from its linguistic form. I will explain the notion of judgement candidate in section 4. Here, I introduce the notion as the successor of the traditional notion of proposition. The traditional, Boethian proposition is linguistic, and has the structure of a declarative sentence (*snow is white*), whereas the Fregean *Gedanke* is non-linguistic and has the structure of a *that*-clause (*that snow is white*). It is thus rather confusing that Russell introduced the term ‘proposition’ as a translation of the Frege’s term ‘Gedanke’. Given that the Boethian term ‘propositio’ is a translation of the Aristotelian term ‘*prôtasis*’, the proposition may not actually be judged, although it has the right form for being asserted. In the *Posterior Analytics*, 72a11–14, Aristotle presents a *prôtasis* as either side of a contradictory pair. Each of these pairs have the right form to be judged, but will not be actually judged at the same time, given that they form a contradiction. It is thus that Ockham makes a distinction between the *propositio iudicativa*, a declarative sentence uttered with assertive force, and the *propositio apprehensiva*, the proposition as understood in its full declarative form, put before the mind without any accompa-

nying act of judgement (Ockham, *Quodlibetal Questions*, Quodlibet 5, Questio 6; cf. Nuchelmans 1980, 27, 74 ff). In section 4, I will argue that the judgement candidate can be understood as a proper bearer of truth and falsity without being committed to a Platonic notion of *Gedanke*. If the thesis is defended that the judgement (candidate) is a proper bearer of truth and falsity, the seven questions asked above are in need of an answer.

## 2. *Some problems with Simons' proper bearers of truth*

Peter Simons claims that proper bearers of truth *and* falsity are not only acts of judgement, but also acts of understanding, acts of assumption, and acts of oral linguistic events such as assertions. All these acts fall under the category of *propositions* in Simons' sense of the term. Simons points out that beliefs are not proper bearers of truth and falsity. For him, the proper bearers are *acts* of judgement and occurrent events of producing or understanding a concrete linguistic token. According to Simons, judgemental acts are events, and as such 'their place and time of occurrence are intrinsic to them' (Simons 2008, 9). Not all philosophers assume that mental acts are a special case of events. Kasimir Twardowski, for example, understands the product of the act to be an event. When we speak about a race as an event, we speak about the result of the activity of racing. A battle is thus an event, but not so the activity of fighting. The action creates the event, but is not itself an event. I come back to Twardowski's distinction between actions and products in the next section. By classifying the act of judgement as an event, there is a danger of neglecting the uniqueness of the act of judgement. For, it seems that one can no longer account for the fact that the act of judgement is of a completely different kind than an event like a storm. Not only in the sense that the act of judgement has the characteristic of being intentional, the act of judgement is also not one of a kind with other mental acts. It is *sui generis*, as Brentano and Frege saw. Independent of the question whether an act is an event, it is important for Simons that the act of judgement is characterised by a certain time and place.

Central to Simons' paper 'Absolute Truth in a Changing World' (Simons 2003) is the claim that this position concerning the bearers of truth can be defended together with the thesis that truth is absolute: truth is not relative to time, space or human beings. According to Simons, we have to

make a distinction between the act of judgement and a declarative sentence as type in order to understand that truth is not relative to time or space. He argues that sentence types are often incomplete as bearers of truth. Language uses indexical terms, and this means that a sentence type like 'It rains' may change from true to false depending on the change of weather. The judgemental act, though, cannot be understood independently of its context; the reference to time and space is built into the act by being the thing it is. And insofar as the production of a sentence token has a fully determined context, the token can be understood as a proper bearer of truth, too (Simons 2003, 46). There is no judgemental act *It rains* independent of a context. A concrete judgemental act of the form *It rains* is determined by 'the expressible content' of the words 'it rains' together with the contextual elements essential to the act (Simons 2008, 11). According to Simons, a particular use of the declarative sentence *is* a proper bearer of truth and falsity, for the utterance is essentially contextual.

This account of truth-bearers is not unlike Twardowski's position in his paper 'On So-Called Relative Truths' (Twardowski 1900). Twardowski argues that if we take sentences as the bearers of truth and falsity, we may be misled in thinking that truth is relative to time, for a sentence with a time indexical may change from true to false. It is therefore the act of judgement that is the proper bearer of truth and falsity, Twardowski says in this paper. Although the act of judgement is dependent upon a human mind, and in space and time, truth is not relative. If a judgement is true, there is nothing that can turn it into a falsehood. And if it is false, there is nothing that can turn it into a truth. For its truth or falsity is determined by the state of things at the time of judging. Without the act of judgement, there is no truth, and as these acts are understood as being in time and space, how can we explain that we can speak of truths no one actually thinks of?

There is an important difference between the position of Simons and that of Twardowski. Twardowski eliminates indexical elements from the act of judgement in order to account for the fact that truth is not relative to time. If we use a declarative sentence to express fully the act of judgement, non-indexical names of places and times are substituted for indexical terms. For Simons, though, indexicality is part of the act judgement. The judgemental act that it is raining, which happens at a certain time and place, is true if and only if it is raining at the time and place of the act of judgement. The truth-conditions for most of our judgements are thus given in contextual terms. Truth itself, though, applies absolutely, and not

relative to a time and place; the contextual elements are essential to the act of judgement, not to its truth-value.

In order to account for the distinction between asserted and unasserted propositions, Simons has to acknowledge besides acts of judgement, acts of understanding as proper bearers of truth. Understanding is here not taken in the sense in which it involves a judgemental act, but in the sense of a mere act of entertaining a thought. If I understand your judgement that John is the murderer, I need not make the judgement myself, but I do apprehend what you are judging. According to Simons, this act of understanding may be true, or false, depending on the state of things at the time of the act. Like the act of judging, the act of understanding is a proper bearer of truth or falsity. We are thus able to speak about the truth of unasserted propositions, where for Simons acts of understanding fall under the concept of *proposition* (Simons 2003, 49). For Simons, the term 'proposition' is the generic term for all proper bearers of truth and falsity. Acts of assumption, acts of assertion and an act of uttering a sentence which is part of a conditional are likewise proper bearers of truth and falsity. There are thus many kinds of things that are properly called true or false. On what basis can Simons call them all 'propositions'? What is it that these acts have in common, besides their being *proper* bearers of truth and falsity? As Simons does not acknowledge propositions or thoughts in Frege's sense, he cannot say that these different kinds of acts all have a Fregean proposition as their content. Neither can he say that these acts have in common that they are the bearers of truth and falsity, for we can call all kinds of things true or false. And saying that they are those acts who are the *proper* bearers of truth and falsity does not explain that there is something special to them that makes them the proper bearers of truth and falsity. What makes these acts of judgement, acts of understanding and acts of assumption unique? Perhaps we can say that they all have a content that may be expressed by a declarative sentence together with the relevant contextual elements rightly fitted in. Is Simons willing to acknowledge such a propositionally structured entity?

Furthermore, not all acts of understanding are proper bearers of truth and falsity. An act of understanding the word 'not' is not true or false. What is, according to Simons, an act of understanding to which truth or falsity pertains? Can we say that it is an act in which we understand a declarative sentence? Simply understanding the words 'It rains' does not constitute, though, a proper bearer of truth, unless it is clear what the context of the token 'It rains' is. The kind of entity that the relevant acts

of understanding, assuming and judging have in common is, I propose, a declarative sentence together with its meaning or expressible content to which indexical elements (time, place, speaker) can be added when appropriate. If one prefers, one may abstract the linguistic expression from this notion, so that the same meaning can be expressed by different linguistic forms. I do not know whether Simons agrees that he is committed to such a notion, but it seems an interesting notion nonetheless, and I will come back to it in the fourth section, where I will argue that this notion need not be identified with the notion of Fregean *Gedanke*. It is rather closer to the notion of proposition introduced by Boethius, and used in the Scholastic tradition, the traditional notion of proposition that may or may not be uttered with assertive force. Finally, the question will be raised whether this traditional notion of proposition is a proper bearer of truth and falsity itself.

It is also important to understand in what sense we might say that an act of understanding is true. When Simons speaks about a true act of understanding, he cannot have in mind an act of understanding whose content is true. The truth of the act of understanding is for him not derived from the truth of the content. So, a true act of understanding must mean something else. The word 'true' seems to have here a special function, that is to be explained by starting with an explanation of the phrase 'false act of understanding'. The word 'false' in this phrase is used as a modifying term, just as the word 'false' in 'false friend,' or 'false gold.' The term 'false' modifies the meaning of 'gold'. If we speak of 'true gold', the term 'true' is a restorative adjective, as Twardowski puts it. It can only be given sense against the background of the notion *false gold*.<sup>1</sup> Neither true gold nor false gold is a special kind of gold. In a similar way, 'false' and 'true' in the phrases 'false act of understanding' and 'true act of understanding' are used as, respectively, modifying and restorative terms. The term 'false' changes the meaning of the term 'act of understanding'. This means that a false act of understanding is not an act of understanding at all. It is not a true, not a proper act of understanding. A true act of understanding is precisely an act of understanding, and nothing more. This is not the way, though, Simons intends to use these phrases.

Is it possible for Simons to explain the validity of *Modus Ponens*? In the first premise,  $\vdash p \rightarrow q$ ,  $p$  is not judged, and has therefore to be understood

---

1. On modifying and restorative terms, see Schaar 2013; on Twardowski's account, see Twardowski 1923.

as an act of understanding. In the second premise,  $\vdash p$ ,  $p$  is judged, and the bearer of truth or falsity is the act of judgement. How can Simons explain that there is a common  $p$  in these two cases? Simons seems to invoke the notion of *expressible content* in such cases together with the contextual elements of the mental act (Simons 2008, 11). Can one speak about the identity of expressible content in these two premises without invoking an abstract notion such as a proposition in the Fregean or the Boethian sense? A certain form of abstraction is at least needed in order to keep the context of the two premises identical, for in the actual reasoning process the two acts happen at different times. Even if we amend this by understanding *Modus Ponens* as applying to products of our acts, there is still a problem because there are two different kinds of acts involved, which makes the products also essentially different. It does not seem, though, that Simons is unwilling to take expressible contents and sentence types into account (Simons 2003, 49). These notions are abstract, though, and one wonders why these sentence types, together with their expressible contents or meanings and elements from the context, cannot be understood as the proper bearers of truth and falsity.

Is it possible for Simons to explain such locutions as ‘That snow is white is true’? Perhaps, Simons would take the term ‘that snow is white’ as a term referring to acts of understanding that snow is white, acts of judgement that snow is white, acts of assumption that snow is white, etc. ‘That snow is white’ would thus be a plural term, just as the term ‘the family’ in the sentence ‘the family is exhausted’, meaning that each of the members of the family is exhausted. In case of ‘the family’ it is clear that we intend to speak about the members of the family, but in case of ‘that snow is white’ it is not at all clear that we intend to speak of all these different kinds of acts. Would it not be more in accordance with natural language to say that we speak here about an identical abstract structure?

Can Simons explain that the assertion ‘John is the murderer.’ is the answer to the question ‘Is John the murderer?’ without invoking an identical content? According to Simons, acts of asserting and understanding are appropriately related to declarative and interrogative sentences. Such an appropriate relation is based on the linguistic competence of the speaker and hearer (Simons 2003, 49–51). Because these linguistic competences involve being able to understand the meaning of the sentence type ‘John is the murderer’ and of the type ‘Is John the murderer?’ the two sentences are related through these competences. Would it not be possible to give a less psychological explanation of the relation between the two sentences?

If Simons is willing to invoke a notion of identical linguistic meaning or expressible content, or that of a sentence type, together with its meaning and a time and place when appropriate, many of the problems mentioned above will disappear. I will argue in the rest of the paper that it is possible to acknowledge such a notion without being committed to Fregean propositions.

What remains the case is that there are two fundamental problems with Simons' account of the proper bearers of truth. There is a diversity of truth-bearers summarized under the name of 'proposition.' Do they form one kind? This would be problematic for acts of judgement, which are generally taken to be *sui generis*: they form a kind of their own. Second, it is not clear what a true act of understanding means. Does it mean an act of understanding whose content is true? In that case, it seems that the propositional content is the proper bearer of truth, which is contrary to Simons' assumption. Or, is the term used as a restorative term, just as we speak of 'true gold', when there is a suggestion of false gold? This point can also be made with respect to the act of judgement. Either a true act of judgement means an act of judgement whose content is true, or the term 'true' is meant to restore the meaning of the term 'judgement' after a suggestion was made that the 'act of judgement' was somehow improper.

### 3. *Twardowski on actions and products*

In his paper 'Actions and Products' from 1912, Twardowski argues that we need to make a distinction between the act of judgement, which is a psychological notion, and the logical notion of judgement product. Twardowski considers questions regarding the act of judgement to be of a psychological nature: we may ask here how the act of judgement arises from other kinds of acts. Questions concerning the judgement product relate to logical questions of validity. When he distinguishes in his lectures on knowledge (Twardowski 1925) between the theory of knowledge and the theory of cognizing, the former is understood as a part of philosophy, whereas the latter is taken to be part of a psychology of thinking. It is for this reason that a theory of knowledge is a theory of judgement products insofar as they are true (idem, 186, 187). In these lectures Twardowski clearly acknowledges that the judgement product, and not the act of judgement, is the proper bearer of truth and falsity (idem, 198). Is it now possible for Twardowski to explain that we may speak of truths no one is actually thinking of?

We may speak of certain beliefs that have persisted through the ages insofar as we have “dispositions owing to which the same sorts of products can arise in us in the future as have in the past.” (idem, 116). We can even speak of the “potential existence” of a judgement insofar as there is a capacity in us to pass this judgement, although no one has actually made this judgement (idem, note 36). An interesting notion to which I come back in the next section.

Insofar as the act of judgement is a purely mental function, the judgement product does not endure, according to Twardowski. Ontologically, the judgement product exists only as long as the act of judgement exists. It is possible, though, for the judgemental act to result in an enduring product insofar as the act of judgement is made manifest in language. We may externalize our act of judgement in an act of assertion, of which the psychophysical product may be a declarative sentence written down on paper. According to Twardowski, this sentence is a sign of the mental product; the sentence expresses the judgement made.

If we compare Twardowski’s position with the one presented in Frege’s *Begriffsschrift*, we see that both aim to capture the judgement made by means of signs on paper. In the *Begriffsschrift*, there is a special sign, the judgement stroke, showing that a judgement has been made. The relation between judgemental stroke and judgemental force is of a unique kind, not comparable to the relation between a sentence and its semantic meaning. The judgemental stroke shows, but does not express, that the judgemental content has actually been judged. For Frege, the judgemental stroke shows how the sentence is used, and is therefore a pragmatic rather than a semantic notion. For Twardowski, the sign for the judgement made seems to be the indicative form or mood of the declarative sentence, just as the interrogative mood would express that a question has been asked. This relation between the indicative mood and the fact that a judgement has been made is not distinguished from the relation between a word and its meaning. Twardowski does not understand that the pragmatic force of the utterance of a sentence is not on the same level as the semantic meaning of a sentence.

Because Frege makes a distinction between the judgemental content and the judgement stroke as a sign of pragmatic force, he is able to distinguish between the semantic meaning, the proposition that is acknowledged to be true, and the pragmatic notion of assertive force. The judgemental stroke is a sign that a judgement has been made, which cannot simply be read from the sentence token as such. One must know what use is made of

this particular token. Twardowski does not acknowledge propositions as abstract meaning-entities; instead, he has only the act of judgement and the judgement made, and this makes it hard for him to make a distinction between a semantic and a pragmatic notion. Is it possible to give a theory in which the distinction between action and product is acknowledged, and allow for a distinction between a semantic and a pragmatic notion at the same time, without invoking the notion of a Fregean proposition or *Gedanke*?

Twardowski's notion of meaning thus involves both the semantic and the pragmatic aspect of the judgement. For Twardowski, the primary meaning of the declarative sentence is the judgement (product) that produced the declarative. The judgement made exists only at the moment of the act of judgement, whereas the sentence has a meaning also after the moment of judging. Therefore, he needs to introduce another notion of meaning.

The second meaning of the sentence is also to be understood in terms of the judgement made. It consists of the judgements made, the mental products, which exist potentially in the sentence, the psychophysical product (Twardowski 1912, 125). They potentially exist in the sentence in the sense that the sentence may be the cause of a judgement produced by someone who reads the sentence. The sentence is able to elicit many mental products in this way. As a good Aristotelian, Twardowski makes a distinction between actuality and potentiality, where the actual is prior in the order of explanation to the potential. Potential existence is not existence in a possible world, it is rather a form of existence in something actual, here a sentence which is the product of a psychophysical act. In the right circumstances, the sentence is the cause of a judgement made, which is similar to the judgement that had this sentence as its effect.

Finally, we can speak about the meaning of a sentence in a third sense insofar as we can distinguish a group of common attributes in the mental products elicited by a particular sentence (*idem*, 127). By an act of abstraction performed on these concrete mental products we are able to speak about an identical meaning, although ontologically speaking there is not an identical abstract meaning. Again, this meaning includes the pragmatic force, as the common attributes need to correspond to the intention with which the sentence was uttered, here, that it was uttered with assertive force (*idem*).

Finally, Twardowski introduces at the end of his paper the notion of *represented judgement*. Is it possible to use this notion as a semantic notion without invoking pragmatic force? In order to explain this notion,

Twardowski introduces the distinction between an artificial and a non-artificial product (idem, 129). The declarative sentence is a non-artificial product insofar as it is the result of an act of assertion or an act of judgement. In standard cases the sentence expresses an assertion or judgement made. On stage, the actions result in artificial products. The actor's sayings and doings do not express real feelings and actions; instead, they express imagined feelings and actions. Declarative sentences on stage are the result of an act in which a judgement is imagined or presented. In general, sentences on stage and other deviant contexts are the result of presentations of assertions, questions and orders, not of real orders or real assertions. Declarative sentences on stage express, according to Twardowski, not a judgement made, but the presentation of a judgement. A presentation of a judgement is not an actual judgement; one thinks of a judgement, but does not make the judgement oneself. The presentations of judgements are also of importance for logic, for the logician often does not judge the examples of premises and conclusions given (idem, 129, 130).

Twardowski is now able to give an account of expressions like 'That snow is white is true.' He could say that the *that*-clause expresses a concept, a product of an act of presentation, which has the possible judgement that snow is white as its object. The account is problematic, though, for we do not attribute truth to a concept when we say that snow is white is true. Perhaps, Twardowski intends to speak about a propositional concept, but that seems to be contradicted by his anti-propositional account of judgement.

Although Twardowski is able to account for the difference between asserted and unasserted proposition by means of his notion of presentation of judgement, the notion is not unproblematic. Twardowski's theory seems not able to account for the fact that sentences on stage and in logic have the same semantic meaning as sentences used in more standard cases. Although it is true that an assertion on stage differs from an assertion in real life, for they differ in pragmatic force, there is also a sense in which they may agree: the sentences used may have the same semantic meaning. Linguistic expressions do not have a different meaning on stage. How is Twardowski able to account for this? For him, sentences on stage and in logic do not express judgements; they express presentations or concepts, whose object is a possible act of judgement. Against this thesis, one may argue that mere presentations or concepts do not have a truth-value. Only the possible act of judgement, which is the object of one's presentation, may be given a truth-value. Presentations and concepts are not of the

same kind as judgements: we cannot evaluate presentations and concepts as true or false.

#### 4. *The judgement candidate*<sup>2</sup>

When someone has danced a dance, we can follow Twardowski and speak of the act of dancing and its product, the danced dance, but there is also a third kind of dance we can talk about. This third meaning of ‘dance’ refers to a dance that may not have been danced yet, although it can be actualized by a dancer in an act of dancing. In a talk presented at a workshop on actions and products, where Peter Simons was present, I called the dance that may not have been danced yet ‘an undanced dance’ and ‘a choreography’.<sup>3</sup> I got some interesting reactions from the public. According to Friederike Moltmann, the idea is problematic because a choreography is itself a product written by a choreographer. According to Jerrold Levinson, the term ‘undanced dance’ is ambiguous. If I understand him well, it may either mean an abstract structure that need not be apprehended by anyone, or it may mean a choreography created at a certain moment in time that may not have been danced yet. In the discussion, Borges’ story of Pierre Menard was mentioned. A written book is like a choreography. Each reader of the book can be compared to a dancer who dances a choreography written by the author. In the Borges story, Menard aims to write a book that is indistinguishable from the *Don Quixote*, not as a copy of Cervantes’ work—he has only vague memories of it—, but as an original work. If the book of Menard would be a mere copy, we could say that he is a reader of the book, who is making a copy at the same time. Menard, though, aims at a completely different book, because it is written at a different time with different literary conventions, although Menard’s book should be type identical to Cervantes’ work. To explain the story, we need one identical, repeatable structure, which allows for two books, two products, like two choreographies, one created in the seventeenth century, and the other at the beginning of the twentieth century.

A dance that may not have been danced yet is, in its first meaning, an abstract structure outside space and time. In its second meaning, it is to be understood as a choreography, created by a choreographer. The cho-

---

2. I called this the assertion-candidate in Schaar (2007).

3. The workshop was held at the IHPST / CNRS, Paris, the 24<sup>th</sup> of January, 2014.

reography itself is an abstract structure, too, but one created at a certain time. The choreography is not itself a dance, but it can be actualised many times in dances danced at different times and places. Important for my argument below is the dance as a choreography created by the dancer or choreographer at a certain time.

Without the notion of dancing, one can neither understand the notion of a dance as abstract structure outside space and time, nor the notion of choreography. The notion of dancing, the *act of dancing*, is in this sense logically prior both to the abstract structure and to the choreography, but a particular act of dancing is never genetically prior to either the abstract structure or the choreography it actualizes. If we focus on the relation between the act of dancing and the choreography, the moment of creation of the choreography may be identical to the moment in which the dance is danced for the first time. Neither is thus genetically prior to the other. The choreography may also be genetically prior to the act of dancing, when it is written down before the actual dance, or just exists hidden in the mind of a choreographer, only known to him. The choreography may never be actually danced, just as a composition may never be actually played. Still, the choreography is there in the sense that it can be actualised in the right circumstances. There is a moment in time in which the choreography has its origin; the moment it is created by the choreographer or the dancer himself, who is here identical to the choreographer. Unlike the abstract structure, the choreography is not an entity existing in a Platonic world waiting to be exemplified. The choreography is abstract, though, in the sense that it is repeatable: different dances at different times and places may actualise the same choreography. If one takes it to be an abstract entity, it is a *created* abstract entity. In a similar way it is possible to speak about a building not yet built, presented on paper or in the head of the architect.

Do we need to acknowledge the dance as abstract structure outside space and time? Are we in need of such a Platonic entity. Without it, we cannot explain that the book Menard aims to write is in some sense identical to the book written by Cervantes. One may answer that Menard has not created a book. Today, his work would simply be understood as plagiarism, not as a creation. It is after all difficult to distinguish the abstract structure outside space and time from Cervantes' product, and some may not want to acknowledge such an abstract structure. Anyway, one may decide not to reify the idea of an abstract structure, and speak of different levels of capacities instead, as is done by Aristotle in the *De*

*Anima* (417a21–417b1). A human being has the capacity to dance, and has thereby a capacity to create choreographies. In this way we may speak of a dance in the first sense. A particular choreography may be written down, and one has thereby created a dance in a second sense, or it may exist in the head of the choreographer, which makes him able to let the dancers dance this dance according to his directions. When the dance is actualised in an act of dancing, the third sense, it results in a danced dance, the product in Twardowski's sense, the fourth sense. And if one thereby knows how to dance that particular dance, one has obtained what Aristotle called a *hexis*, a *habitus*, as it is translated by the Scholastics. One has obtained a potentiality to dance that particular dance.

If we apply these distinctions to the notion of judgement, we get five meanings of the term 'judgement'. When human beings have learnt a language, and have learnt the game of giving and asking for reasons, they have *the capacity to judge*, the first meaning of 'judgement'. Thereby, one has the capacity to express or apprehend what I will call a *judgement candidate*, the second meaning of 'judgement', to be explained in this section. Let us assume that no one had thought of Goldbach's conjecture before Goldbach did. When Goldbach formulated his conjecture, he created a judgement candidate, a judgement that no one thought of before, but had already been there in our capacities to judge (or as an abstract entity outside space and time, if one prefers the Platonic terminology). Goldbach created the judgement candidate by formulating his conjecture that every even integer greater than 2 can be expressed as the sum of two primes. The judgement candidate is in this sense on a par with a choreography created by a choreographer. Once, there might come a time when Goldbach's conjecture is demonstrated: it would thereby be actualised in an *act of judgement*, the third meaning, resulting in a *judgement made*, the fourth meaning of the term 'judgement'. Finally, we may say that by having made this judgement once, we have obtained a potentiality to make that particular judgement, which potentiality is sometimes called a *belief*, the fifth meaning of the term 'judgement'.

What is the judgement candidate? The notion differs from the Fregean notion of *Gedanke*, and is closer to Boethius' notion of proposition or the Scholastic notion of proposition that may either be merely apprehended or be actually judged. The judgement candidate may be understood as relatively independent of the declarative sentence, but it may also be understood as a declarative sentence type together with its meaning. Whereas the *Gedanke* has the form *that snow is white*; the form of the judgement

candidate is *that snow is white is true* or *snow is white*, so it includes the indicative mood of the sentence, in contrast to the corresponding wish candidate *let snow be white* (*let it be true that snow is white*). The judgement candidate has the form of a declarative sentence, but no assertive force is involved. If one has fully apprehended a judgement candidate, it should be clear under what condition one would be entitled to make the judgement. When I am wondering whether I would make the same judgement that someone else is making, I apprehend a judgement candidate. I consider the condition under which I would be entitled to pass the judgement myself. If I am entitled to make the judgement, my judgement is true in the sense of correct. The judgement candidate is thus explained in terms of judgemental correctness conditions (see below).

How do contextual elements enter into the judgement candidate? There is no actual judgement, which would be able to relate the judgement to a certain context. The judgement candidate is not situated in time and space the way the act of judgement is. The judgement candidate that it rains here and now would be incomplete, as Simons saw, if we understood the judgement candidate in purely linguistic terms. One has to understand what time and place is relevant in order to be entitled to make the judgement. The way we refer to the relevant time and place may be simply by means of indexical terms and an apprehension of the relevant context. The judgement candidate is thus not always expressed by linguistic means alone. Apprehending the judgement candidate is in some cases a combination of apprehending the meaning of, for example, 'it rains' together with an apprehension of the relevant time and place. Apprehending the same judgement candidate a day later may involve some reconstruction on our part. The judgement candidate that it is raining today will tomorrow be apprehended by means of the word 'yesterday' and a possibility to apprehend the first day by its relation to the second day, that it is the day before. Perhaps, though, we should say that these sentences express different, though related, judgement candidates. It may sometimes be enough to determine that it is raining now, without knowing what day it is beyond knowing that it is today; tomorrow, it might be demanded to identify the same judgement candidate by means of the apprehension of that day as the day before. The apprehension of contextual elements by non-linguistic means may thus be essential to the apprehension of some judgement candidates. The notion of judgement candidate is different from Simons' notion. For Simons, the time and place of judgement are 'external to the judgement's content' (Simons 2008, 11). In answer to Simons we

may say that when I judge today that it is raining, what I judge is that it is raining on that day, not simply that it is raining. It is in agreement, though, with Simons' thesis that contextual elements may form an essential part of the judgement. Often, we have no means to use non-indexical terms for these contextual elements, while it is perfectly clear which contextual elements are involved in the judgement. A final point, we need not argue that the day, the place and the speaker themselves form part of the judgement candidate, it may rather be something to be called 'the form' of this particular day, place or speaker, without its matter, which can be understood to be part of the candidate notion. Such a form need not be a linguistic expression. It is thus not persons themselves that would be part of the judgement candidate.

What can *correctness* of the judgement mean? In principle, there are two possibilities. One may give either a knowledge or a truth account of assertion, and if judgement is the interiorisation of assertion, one may give an account of the judgement candidate either in epistemic or in non-epistemic terms. The arguments for an epistemic account of assertion and judgement are strong, and I have defended elsewhere that one is entitled to make an assertion or to pass a judgement only if one is able to justify one's assertion or judgement (see Schaar 2011). On this account, *judgemental correctness* is to be explained in terms of justification. The judgement made is *correct* if and only if it *has been justified*. The judgement candidate cannot be called correct in this sense, for the judgement may not have been made yet, and may therefore not be actually justified. The judgement candidate is correct in an extended sense if and only if it is *justifiable*. Judgemental correctness in the extended sense is not subjective in the sense that it depends on a judging agent actually passing and justifying the judgement. There is an internal relation, though, between the two correctness notions. We can now formulate an identity criterion for the judgement candidate: two sentences express the same judgement candidate if and only if one is entitled to utter these sentences with assertive force under the same epistemic conditions.

It is now quite natural to make a distinction between an asserted and an unasserted utterance of a declarative sentence. Both express a judgement candidate, while the former manifests an assertion or judgement made, as well. What is judged and what is merely apprehended is in both cases the judgement candidate. It is thus the same judgement candidate that is asserted or judged in one case, and merely apprehended in the other case. Only in the former case, a pragmatic force is present. In this way we

can distinguish between the semantic notion of the judgement candidate and the notion of judgement or assertion involving pragmatic force. If we want to make the pragmatic force explicit in our system, we can use Frege's judgemental stroke or Russell's assertion sign. Notwithstanding the fact that the judgement candidate is a semantic notion, it is explained in terms of the pragmatic notion of judgemental act, or assertive force.

If someone utters a conditional sentence with assertive force, the antecedent is uttered without assertive force, and expresses a mere judgement candidate. The same judgement candidate may also be expressed by an utterance with assertive force. It is thus possible to explain *Modus Ponens*, without invoking a Platonic notion of *Gedanke*.

With the notion of judgement candidate we are entitled to speak of truths not actually thought of, but we do not seem to be entitled to speak of truths that have never been thought of. As soon as Goldbach's conjecture will be demonstrated, we are entitled to say that the judgement candidate always had been correct in the extended sense. This does not imply, though, that we are thereby entitled to speak about a Platonic notion outside space and time that has always been correct independent of anyone formulating the judgement candidate for the first time.

In what sense can it be said that a judgement is incorrect? A judgement may be said to be incorrect, insofar as the judging agent is not able to justify his judgement. This seems to make the incorrectness of the judgement to be a subjective notion. There is another sense of incorrectness that may be applicable to the judgement candidate. We may *suppose* that the judgement candidate is justifiable, and if the supposed correctness of this judgement turns out to be inconsistent with other judgements already justified, that is, if these judgements cannot all be correct, we are entitled to say that the judgement candidate is incorrect in the sense that it cannot be justified. The correctness and incorrectness of the judgement is thus a human affair, and therefore does not imply infallibility. As we may be mistaken, judgemental correctness is not the only role of truth (cf. Sundholm 2004).

Does this account of truth-bearers imply a form of subjectivism? On the one hand, there is no correctness of the judgement independent of the possibility of a judging agent justifying the judgement. And the notion of judging agent is used in the explanation of the judgement candidate. Furthermore, judgemental correctness in the strict sense is dependent upon the agent's actually having justified the judgement. On the other hand, judgemental correctness in the extended sense is independent of an agent who actually passes the judgement. Furthermore, the explanation of the

judgement candidate is not subjective. For, what counts as justification for a judgement is not a private matter, nor is it something that can arbitrarily be decided. It is given by the explanation of the judgement candidate.

Is it possible to relate judgements to questions, orders and promises? How can we explain that a certain judgement is the answer to a question, that it gives the satisfaction of a demand or the fulfilment of a promise? We may explain the question candidate in terms of its answerhood conditions, and thus in terms of the judgement candidates that would count as an answer to the question. Likewise, we can explain the demand candidate in terms of its satisfaction conditions, and the promise candidate in terms of its fulfilment conditions. All three are thus explained in terms of the judgement candidate. There is thus an essential relation between these notions and the notion of judgement candidate.

Finally, we need to answer the question what the that-clause in 'That snow is white is true' refers to. Equally, we have to give an account of the that-clause in sentences like 'You promised me that you would come', 'I asked you that you would hand in the paper by Friday'. Such expressions can be given a more natural form: 'You promised me to come', 'I asked you to hand in the paper by Friday.' Sometimes, we cannot use a that-clause, as in 'I asked him what the weather would be like for tomorrow.' If we want to give a general enough explanation for these clauses, we also have to take the last example into account. What is referred to by the clause 'what the weather would be like for tomorrow' cannot be a Fregean *Gedanke*. For, *what the weather would be like tomorrow* has no truth-value. Does the clause refer to a question made or a potential question? We also have to give an account of the meaning of the sentence 'I didn't ask him what the weather would be like'. If we want to give a unified account, the question made is thus excluded, and we may conclude that a question candidate is referred to by such a clause. In a similar way, we can say that the that-clause of 'He asserted that snow is white' refers to a judgement candidate, which has the role here of judgemental content. From there, one may conclude that the that-clause in 'That snow is white is true' refers to a judgement candidate, too. The answer is not so straight-forward, though, if one acknowledges both Fregean *Gedanken* and judgement candidates. The answer depends on an analogy between the different examples, and one may argue that different uses of the that-clause are involved.

Finally, as the proper bearer of truth and falsity is not in space and time, we have a guarantee that truth and falsity are not relative to space and time. The fact that some judgement candidates cannot be apprehended

without apprehending the relevant time and place does not mean that the judgement candidates themselves are in time and space in the way tables and chairs are in time and space. While defending that the judgement (candidate) is the proper bearer of judgemental correctness, we have answered the questions addressed in the introduction, without invoking a Platonic notion of proposition.

#### REFERENCES

- Nuchelmans, Gabriel 1980: *Late-Scholastic and Humanist Theories of the Proposition*. Amsterdam, Oxford, New York: North Holland Publishing Company.
- Schaar, Maria van der 2007: "The Assertion-candidate and the Meaning of Mood". *Synthese* 159, 61–82.
- 2011: "Assertion and Grounding: A Theory of Assertion for Constructive Type Theory". *Synthese* 183, 187–210.
- 2013: "Wooden Horses and False Friends; on the Logic of Adjectives". In: Kevin Mulligan, Katarzyna Kijania-Placek, & Tomasz Placek (eds.), *The History and Philosophy of Polish Logic; Essays in Honour of Jan Wolenski*. Basingstoke: Palgrave Macmillan, 103–116.
- Simons, Peter 2003: "Absolute Truth in a Changing World". In: Jaako Hintikka, et al. (eds.), *Philosophy and Logic in Search of the Polish Tradition; Essays in Honour of Jan Wolenski*. Dordrecht: Kluwer, 37–54.
- 2008: "Twardowski on Truth". *The Baltic International Yearbook of Cognition, Logic and Communication* 4, article 7, 14 pages; available at <http://newprairiepress.org/biyclc/vol4/iss1/7/>
- Sundholm, Göran 2004: "Antirealism and the Roles of Truth". In: I. Niiniluoto, Matti Sintonen & Jan Wolenski (eds.), *Handbook of Epistemology*. Dordrecht: Kluwer, 437–466.
- Twardowski, Kazimierz 1900: "On So-Called Relative Truths". In: Twardowski 1999, 147–169.
- 1912: "Actions and Products. Some Remarks from the Borderline of Psychology, Grammar and Logic". In: Twardowski 1999, 103–132.
- 1923: "On the Logic of Adjectives". In: Twardowski 1999, 141–143.
- 1925: "Theory of Knowledge. A Lecture Course". In: Twardowski 1999, 181–239.
- 1999: *On Actions, Products and Other Topics in Philosophy*. Johannes Brandl & Jan Wolenski (eds.), Amsterdam: Rodopi.



PART III  
LOGIC



## BOLZANO'S DEFINITION OF ANALYTIC PROPOSITIONS

Bob HALE  
University of Sheffield  
Crispin WRIGHT\*  
New York University

### *Summary*

We begin by drawing attention to some drawbacks of what we shall call the Frege-Quine definition of analytic truth. With this we contrast the definition of analytic propositions given by Bolzano in his *Wissenschaftslehre*. If Bolzano's definition is viewed, as Bolzano himself almost certainly did not view it, as attempting to capture the notion of analyticity as truth-in-virtue-of-meaning which occupied centre stage during the first half of the last century and which, Quine's influential assault on it notwithstanding, continues to attract philosophical attention, it runs into some very serious problems. We argue that Bolzano's central idea can, nevertheless, be used as the basis of a new definition which avoids these problems and possesses definite advantages over the Frege-Quine approach. Our title notwithstanding, we make no claim to contribute to the exegesis of Bolzano's thought and works, which we must leave to those more expert in these matters than we are. Naturally, we have done our best not to misrepresent Bolzano's views, and believe we have avoided doing so. But it bears emphasis that it is no part of our intention to suggest that the modifications to his definition which we propose would have had any appeal for him, or that he had, or would have had, any sympathy with the project which motivates them.

### 1. *Frege's definition*

A noteworthy feature of Frege's explanation of the distinction between analytic and synthetic judgements is that he views the distinction as an

\* We are both pleased to be able to contribute to this special issue for Peter, and grateful to Sandra Lapointe for inviting us to do so. In addition to his worthy contributions to the philosophy of mathematics and metaphysics, Peter has made a huge contribution to our appreciation and understanding of Central European philosophy and logic. It is, accordingly, an added pleasure to contribute a paper on a topic close to his intellectual heart.

epistemological one, in parallel with the obviously epistemological distinction between a priori and a posteriori judgements:

Now these distinctions between a priori and a posteriori, synthetic and analytic, concern ... not the *content* of the judgement but the *justification* for making the judgement. ... When a proposition is called a posteriori or analytic, in my sense, this is not a judgement about the conditions, psychological, physiological and physical, which have made it possible to form the content of the proposition in our consciousness; nor is it a judgement about the way in which some other man has come, perhaps erroneously, to believe it true; rather, it is a judgement about the ultimate ground upon which rests the justification for holding it to be true.(Frege (1974), §3)

Clearly Frege's main concern here is to distance himself from any sort of psychological account of the distinctions he is about to explain, and from any suggestion that they relate to different ways in which judgements are to be causally explained. But it is worth emphasizing that in holding that the distinctions concern justification, he is also distancing himself from, or at least avoiding commitment to, any view on which the distinctions concern the grounds of truth—what *makes* the judgement true—in the way that is suggested by, for example, subsequent characterizations of analyticity in terms of 'truth-in-virtue-of-meaning'. We shall return to this point, and its significance, much later. Frege continues:

This means that the question is removed from the sphere of psychology, and assigned, if the truth concerned is a mathematical one, to the sphere of mathematics. It now becomes a problem of finding the proof of the proposition, and of following it back to the primitive truths. If in the course of doing so, we come only only general logical laws and definitions, then the truth is an analytic one, bearing in mind that we must take account also of any propositions on which the admissibility of any definition depends.

Thus according to Frege, a judgement is analytic iff the proposition judged true can be proved from using only general logical laws, together with definitions.

There is an obvious similarity between Frege's definition and Quine's subsequent characterization of what he terms the 'second class' of statements generally held to be analytic. The 'first class' of such statements are those, such as 'No unmarried man is married', which, he says, 'may be called *logically true*, where a logical truth is a statement which is true and remains true under all reinterpretations of its components other than the logical particles'. But, he continues

... there is also a second class of analytic statements, typified by:

(2) No bachelor is married

The characteristic of such a statement is that it can be turned into a logical truth by putting synonyms for synonyms. (Quine (1953), 22f.)

We might, then, define:

A statement  $S$  is *broadly analytic* iff (i)  $S$  is logically true, or (ii) for some logically true statement  $S^*$ ,  $S$  is transformable into  $S^*$  by substituting synonymous expressions.

Statements which qualify as broadly analytic by clause (i) may be said to be *narrowly analytic*. Although, for reasons too well-known to require restatement here, Quine himself does not regard this as an acceptable definition, no harm need result from labelling it as the *Frege-Quine definition* (of broad analyticity).

The Frege-Quine definition has two notable drawbacks. The first concerns logical truths. Such truths compose the base class in terms of which the remainder of the class of broadly analytic truths is defined. But while it is clear that statements in the remainder are supposed to count as analytic because reducible to logical truths, the status of logical truths themselves as analytic is left entirely without explanation. The point is not that the choice of logical truths to compose the base class is arbitrary—just about anyone who has any use for the notion of analyticity would classify them as analytic. And everyone would agree that it would be absurd to take instead, say, the laws of thermodynamics, or the truths recorded in Mrs. Beaton's *Manual of Cookery and Household Management*, as the base class. The point is just that the definition gives no hint why logical truths should themselves be regarded as analytic.<sup>1</sup>

A second drawback concerns the extensional correctness of the definition. If we think of it, not as a straightforward stipulation, but as intended to codify an already accepted notion, then it seems clearly to fail. For there appears to be a significant class of statements which those who think

---

1. We are not suggesting that this drawback is one which Quine would, or should, have worried about. It is a drawback only for someone who is trying, as Quine was not, to give an acceptable definition which does not merely circumscribe the extension of the term 'analytic' but captures the essence of the concept. Quine thought there was no essence to capture, and was merely trying to characterize, for critical purposes, the class of statements commonly taken to be analytic. Whether he should have been worried by the second shortcoming to which we draw attention is another question entirely.

they understand the notion would wish to see classified as ‘true-purely-in-virtue-of-meaning’, so as analytic in the intended spirit of the notion, whose members are neither logical truths nor reducible to logical truths by substitution of synonyms for synonyms. Well known candidates are such statements as ‘Anything red is coloured’, ‘If one event precedes another, and the second precedes a third, then the first precedes the third’—the reader will surely be able to think of many others. Perhaps some candidates are more controversial than others—witness ‘Nothing can be red and green all over’—but that there is a substantial class of statements which fall under the intuitive extension of ‘analytic’, yet elude classification as analytic by the Frege-Quine definition because they essentially involve terms which do not admit of the definitional paraphrases which would permit their reduction to logical truths, seems beyond serious question.<sup>2</sup>

## 2. Bolzano’s definition

In his *Wissenschaftslehre* (Bolzano (1837)), volume II, section 148, Bernard Bolzano gives a definition of analytic propositions which holds out some promise of addressing the last point. His definition of what he calls narrowly or logically analytic propositions is of some historical interest, because it anticipates by about 100 years the definition of logical truth given by Quine mentioned above.<sup>3</sup> Bolzano takes being true and being false, being analytic and being synthetic to be properties, in the primary sense, of what he calls *propositions in themselves* [Sätze an sich], which he distinguishes both from verbal and mental propositions. He takes propositions to be structured entities composed of ideas or concepts. In this chapter, he considers the effects of varying some of the ideas that make up a proposition, whilst keeping the other ideas involved in it fixed. What he means by varying an idea here is replacing it uniformly throughout a proposition by another idea. He notices that some propositions are such that if we keep only the logical ideas or concepts occurring in them fixed,

---

2. The second of these drawbacks, and something close to the first, are pointed out by Paul Boghossian in a couple of places (see Boghossian (1997), 338f., Boghossian (1996), 368).

3. As Quine acknowledges—see Quine (1966), fn.2, 103. In saying the Bolzano’s definition anticipates Quine’s, we are claiming only that the central idea of Quine’s definition is already present in Bolzano’s, not that the two are equivalent. They are not. Most importantly, Quine’s has the unfortunate consequence that such sentences as  $\exists x \exists y x \neq y$  qualify as logical truths, whereas they are not logically analytic by Bolzano’s.

we may vary any of the remaining ideas without changing the truth-value of the proposition.<sup>4</sup> It is these propositions which he defines to be logically analytic, or analytic in the narrower sense. Schematic examples he actually gives are: 'A is A', 'An A which is a B is an A', 'An A which is a B is a B', and 'Every object is either B or non-B'.

If we say, in accordance with a well-established terminology, that an expression occurs *essentially* in a statement if and only if uniformly replacing it throughout that statement may result in a statement that differs in truth-value from the original one, and give a parallel explanation of an idea's occurring essentially in a proposition, then we can see that Bolzano's definition of logical analyticity is virtually the same as Quine's definition of logical truth: for Bolzano, a proposition is logically analytically true if and only if it is true and only logical ideas or concepts occur in it essentially; while for Quine, a statement is logically true iff it is true and contains only logical expressions essentially. The interest of Bolzano's definition is not, however, confined to its being a forerunner of Quine's. Like Quine, Bolzano makes a distinction between broader and narrower analytic truths. But whereas for Quine the broader notion is to be explained, if it can be explained at all, on the basis of the narrower one, Bolzano reverses the direction of explanation—for him, it is the broader notion which is basic, and logically or narrowly analytic truth is merely a special case of it.<sup>5</sup> To understand how this comes about, we need to look more closely at his explanation.

Bolzano's general concern (see especially §147) is with the effects of varying one or other of the ideas in a proposition on its truth-value. Let  $p$  be any proposition, and let  $i_1, \dots, i_n$  be the ideas of which it is composed.

---

4. More precisely, the result of varying these ideas will be a proposition having the same truth-value, if it has *denotation* at all. By saying that an idea is *denotative*, Bolzano means that it 'has an object falling under it' (see Berg (1973), 82). Bolzano's word is *gegenständlich*). In the case of propositions, the result of substituting of one idea for another may be a proposition which fails to have the same truth-value because it lacks denotation altogether. Bolzano gives the example 'The man Caius is mortal', telling us that while every replacement for the idea of Caius must yield a true proposition if it yields a proposition with denotation at all, it may be that an idea is substituted—such as the idea of a rose or a triangle—which results in a proposition lacking denotation altogether. See Berg (1973), 188f.

5. It is, of course, no accident that Quine privileges the narrow notion. For he believes that while the broader notion cannot be satisfactorily explained, because an explanation requires appeal to the problematic notion of synonymy, or some equally problematic alternative, the narrow notion can be explained, drawing only upon the unproblematic notions of truth and uniform substitution. Whether he is right so to believe is not our concern here. For an early statement of the case against, see Strawson (1957).

Take one of these ideas,  $i_k$ . Then in general, some of the results of varying  $i_k$  by putting another idea in its place will be true propositions, and some will be false. Roughly speaking, Bolzano defines the degree of validity<sup>6</sup> of  $p$  with respect to  $i_k$  to be the ratio of true propositions that result from varying  $i_k$  to the total number of propositions that are obtainable by varying  $i_k$ . In the limiting case when every proposition that so results is true—so that the validity of  $p$  with respect to  $i_k = 1$ —Bolzano says that the proposition is universally valid with respect to  $i_k$  (or universally invalid, if every resulting proposition is false). We could express this by saying that the idea  $i_k$  occurs inessentially in  $p$ . Bolzano then, in effect, defines a proposition to be analytic, in his broad sense, if it contains at least one idea inessentially.

This proposal contrasts with Frege's, and with the Frege-Quine definition, in several respects.

First, whereas Frege and Frege-Quine seek to define analytic *truth*, Bolzano's definiens is analyticity—i.e. *analytic truth-or-falsehood*. For him, analytically *true* and analytically *false* propositions are simply propositions which are both analytic and true, or analytic and false, respectively. This, as we shall see in due course, is a source of some difficulty; but for now, we simply note the point.<sup>7</sup>

Second, the definitions diverge over the *bearers* of analyticity. For Frege, analyticity is a property of *judgements*, and for Frege-Quine, of *statements*, while for Bolzano, it is a property of *propositions-in-themselves*. This difference may be of some significance for the detailed exegesis of Bolzano's own view, but that is not our business here. It is straightforward enough to transpose Bolzano's definition so as to apply to statements, and while we shall respect his usage when reporting or commenting on his actual views, we shall switch, without special notice, to taking statements as the bearers of analyticity, when we come to consider modifications of his definition.

There is a third, far more important point of contrast, at least between Bolzano's definition and Frege's: while, as we have noted, Frege takes being

---

6. Bolzano's term is *Gültigkeit*, which Rolf George (George (1972)) translates as 'satisfiability'; Jan Berg's translation (Berg (1973), 187) has 'validity'; 'degree of validity', which seems to us more accurate, was suggested by Wolfgang Künne. Our formulation omits some restrictions Bolzano introduces, but which do not affect our discussion.

7. Curiously, Bolzano's examples of logically analytic propositions are all examples of true propositions; but he does give as examples of analytic propositions some which he clearly takes to be false. See e.g. Berg (1973), 192, where he cites 'A morally evil man nevertheless enjoys eternal happiness' as an analytic proposition which remains false under any substitution for the idea of *man*.

analytic, like being a priori or a posteriori, to be fundamentally an *epistemological* notion, there is no whiff of epistemology in Bolzano's account of it. His concern is simply with the effects of varying certain of the ideas composing a proposition upon its truth-value. Once again, this is a point to which we shall return in the sequel.

### 3. *Potential advantages of Bolzano's definition*

Well and good. The question arising now is what, if any, may be the advantages of Bolzano's definition over that of Frege-Quine.<sup>8</sup> One apparent such advantage may speedily be seen to be illusory. It is clear that logically analytic propositions are, for Bolzano, a special case of analytic propositions in his broader sense. For logically analytic truths will be true propositions in which *all* but logical ideas occur inessentially. It may now appear that the primary advantage of Bolzano's definition is that it captures a broader notion of analytic truth, corresponding to Quine's second class, whilst deploying only the relatively modest resources—viz. the notions of truth and uniform substitution—which Quine thinks sufficient to characterize the narrower class of logical truths. It may thus appear that Bolzano provides a way of bypassing the difficulties Quine raises about the explanation of the broader notion—that he succeeds in defining it without reliance upon the notion of synonymy or any of the other notions Quine regards as equally suspect. It is important to see that this apparent advantage is merely apparent.

The reason why this is so becomes clear as soon as we ask whether the proposition expressed by, for example, 'Vixens are female' qualifies as (broadly) analytic in Bolzano's sense. At first sight, it fails to do so, since—to put the difficulty in Bolzano's terminology—it appears to contain no idea that can be varied at will without variation in truth-value. If the written proposition is assumed accurately to reflect the composition out of ideas of the proposition in itself that the sentence 'Vixens are female'

---

8. Our focus here is entirely on the potential advantages of Bolzano's definition, when it is viewed as an alternative to the more familiar Frege-Quine definition. As we say in our abstract, we make no claim concerning what may have been Bolzano's own purposes in defining analyticity, what role his definition may have been intended to play in his overall philosophy, or what relation he may have taken it to bear to Kant's definition(s). For interesting discussions of these and other questions about Bolzano's actual views, see Künne (2008a), reprinted in Künne (2008b), which contains several other relevant papers, and Lapointe (2011), chs. 4, 5.

expresses, then that proposition must be reckoned synthetic in Bolzano's sense; for it will contain no idea inessentially. And so it will be with indefinitely many further examples of propositions which would be classified as analytic, at least by anyone who has any use for the (broader) notion at all.

Of course, Bolzano would regard the proposition expressed by 'Vixens are female' as analytic, even though it appears at first to fail to qualify as such by his definition. In a note on his definition, he says:

In order to determine whether a proposition which is given a certain linguistic expression is analytic or synthetic, more is required than a cursory inspection of its words. A proposition may be analytic, perhaps logically analytic, or even identical, though its literal phrasing does not make this immediately apparent. ... Thus it may not be immediately obvious that the proposition 'Every effect has a cause' is in fact identical, or at any rate analytic; for by 'effect' we always mean something which is brought about by something else, and the phrase 'to have a cause' means as much as 'to be brought about by something else'; thus the above proposition merely means 'Whatever is brought about by something else is brought about by something else'<sup>9</sup>

If we say, as Bolzano would presumably have been happy to say, that a spoken proposition is analytic if the proposition-in-itself expressed by it is so, then the point he is making here could be put by saying that the proposition-in-itself (see page 4) that is expressed by a given spoken proposition is that proposition-in-itself that results from the given spoken proposition by fully expanding it accordance with definitions of its ingredient words. But this means, of course, that to justify the acknowledgement of the proposition expressed by 'Vixens are female' as analytic, Bolzano has after all to rely upon claims about synonymy, and so has not after all provided a way of explaining broad analyticity that both gives it the intuitively correct extension and bypasses Quine's objections to the notion.<sup>10</sup>

A genuine advantage of Bolzano's definition, assuming it to be acceptable, lies elsewhere—in its greater generality. Specifically, it promises to accommodate as analytic examples of the 'third kind' which fail to be so classified by the Frege-Quine definition. Putative examples, again, are:

---

9. Bolzano (1837), §148. By an 'identical' proposition Bolzano means an instance of the Law of Identity 'A is A'. A quite different explanation how Bolzano can count the propositions expressed by 'No bachelor is married' and 'Vixens are female' as analytic is suggested by Lapointe (see Lapointe (2011), 64ff.). We see no good reason not to adopt the simpler one suggested in the this passage.

10. The point we have been emphasizing is made very clearly by Wolfgang Künne (see Künne (2008a), 298ff.)

If Mozart's stockings are yellow, then they are coloured.  
If Vivaldi's birthday precedes Händel's, and Händel's precedes Bach's,  
then Vivaldi's precedes Bach's.

For in these propositions, the ideas of Mozart's stockings, and Vivaldi's, Händel's and Bach's birthdays all occur inessentially. And with a small refinement of Bolzano's definition, we can ensure that the more general propositions such as:

Anything yellow is coloured.  
If one event precedes a second, and the second precedes a third,  
the first precedes the third.

also qualify. Of course, they do not qualify on his definition as it is, because neither contains any idea inessentially. But it would not be unreasonable to claim that a generalization is analytic iff all its instances are, and to modify the definition accordingly. Under the modified definition, these and similar general statements *would* qualify. Thus Bolzano's definition together with our modest emendation appears to have the very desirable consequence that just the kind of true statement which we previously claimed ought to count as analytic—but fails to do so on the Frege-Quine definition—gets correctly classified. So although Bolzano's definition does not dispense with reliance on the notion of synonymy, it does allow us to recognize as analytic many statements which are not reducible to logical truths by synonymous substitution.

Further, there is at least some progress with the other drawback of the Frege-Quine definition—the unexplained status of logical truths as analytic. For since, on Bolzano's definition, logically analytic propositions are just a special case of analytic propositions in general, there is no special problem about explaining why they are analytic. But only partial progress—for obviously, assuming the definition to be otherwise in good standing, there would still be a good question why it should be thought to capture whatever intuitive idea informs our application of the notion of analyticity. But before pursuing that question, we should face up to the fact that the definition is open to a seemingly fatal line of objection.<sup>11</sup>

---

11. The objection we are about to consider is, of course, an objection to the definition when it is viewed as attempting to capture the notion at which Frege-Quine is aimed—a notion on which analytically true propositions will be invariably necessary and knowable a priori. This perspective is assumed for the remainder of the paper, and in particular, by our claims about the

#### 4. *Over-extension (1)*

We have taken one of the advantages of Bolzano's definition to consist in its capturing a broader range of analytic truths than the Frege-Quine definition. But if the objection we are going to consider is sound, the definition is too broad—because it has the consequence that many propositions are to be reckoned analytically true that are not so, but are very plainly at best statements of contingent empirical fact.

Consider any contingently true generalization—this can be either some true statement of natural law, or equally some true accidental generalization. For simplicity, and without loss of generality, we may suppose it to have the form:  $\forall x(Fx \rightarrow Gx)$ . Now consider any one of its instances:  $Fa \rightarrow Ga$ . Then under the supposition we are making, this statement is not merely true, but remains so under any uniform substitution on  $a$ . Accordingly, whilst the parent generalization no doubt comes out as synthetic under Bolzano's definition—there being, we may assume, (uniform) substitutions on  $F$  or  $G$  (or both) which yield a falsehood—the instance qualifies as analytic. Thus it is true—though presumably not in consequence of any natural law—that no eighteenth century philosopher died on the anniversary of his birth. Thus whatever substitution is made for 'Kant' in 'Kant was not an eighteenth century philosopher who died on the anniversary of his birth', a true statement results. Hence our proposition about Kant must be reckoned *analytically* true, on Bolzano's account.

This objection, noted by William and Martha Kneale<sup>12</sup>, appears quite devastating—for it appears that the very feature of Bolzano's definition in virtue of which it promises to capture a wider notion than either that of narrow or logical analyticity (i.e. logical truth, as defined by Quine) or broad analyticity as explained in terms of reducibility to logical truth via

---

potential advantages of Bolzano's definition and of the modifications of it we consider. Whether the Kneales' and the other main problem we consider are problems for Bolzano's own project is not our concern here. See also note 15 below.

12. See Kneale (1962), 366f.; the Kneales are also responsible for the nice example. They clearly assume the perspective on Bolzano's definition described in the note 12. A kind of obverse of their example may be got by considering false existential generalizations: if  $\exists xA(x)$  is false, then  $a$  will occur inessentially in  $A(a)$ ; so that each and every instance of the generalization will count as analytic in Bolzano's sense, regardless of the status of the parent existential generalization. Clearly there will be further anomalies. Thus consider and statement  $\forall x(Fx \vee (\neg Fx \wedge p))$ , where  $p$  is some contingent truth. Any instance  $Fa \vee (\neg Fa \wedge p)$  will rank as analytic. Of course, were  $p$  false rather than true,  $a$  might well fail to occur inessentially, since  $Fa$  might be true but  $Fb$  false.  $a$ 's inessential occurrence is contingent on the truth-value of  $p$ : This—contingently inessential occurrence—is what the Kneales' and similar examples exploit.

definitional expansion, is precisely what is responsible for the disaster.<sup>13</sup> If we view Bolzano's definition as an attempt to generalize Quine's definition of logical truth, the generalization amounts to this: whereas Quine requires for a statement to be logically true that *all* the non-logical expressions occurring in it should do so inessentially, Bolzano requires (for a statement to be analytic) only that *some* of the non-logical expressions occurring in it should do so. But in any instance of a true general statement, the singular terms will occur inessentially, so that any such statement will count as analytic. The resulting unwanted expansion of the class of analytic truths thus appears as the inevitable, and clearly unacceptable, price of seeking to define a broader notion in terms of the incapacity of uniform substitution to change truth value.

Is there any way to meet this difficulty? Can we find a revision of Bolzano's definition which retains its advantages whilst avoiding this consequence?

## 5. *Blind alleys*

### 5.1 *A two-part definition?*

Whilst Bolzano's definition misclassifies as analytically true any instance of a synthetically true generalization, it appears to yield the right verdict when applied to the parent generalization itself. Thus there is, for example, no idea for which we may freely substitute any other idea in the proposition expressed by 'No eighteenth century philosopher died on his birthday'. More generally, Bolzano's definition appears to yield intuitively correct results when its application is restricted to propositions expressed by sentences devoid of singular terms. Thus it may seem that we could secure

13. It appears so, but is it so? In fact, an analogue, or at least a close relative, of the Kneales' problem afflicts Quine's definition of logical truth itself, independently of Bolzano's generalization. As is well-known, for any natural number  $n$  we can express that there exist at least  $n$  objects in the language of first-order quantification theory with identity, for example by writing  $\forall x \exists_{n-1} y y \neq x$  (where  $\exists_n x$ , meaning 'There are at least  $n$   $x$ ', is recursively definable in the usual way). A nominalist who thinks that there are only concrete objects, but that there are at least 17 of them, will take  $\forall x \exists_{16} y y \neq x$  to be true, but it is surely not a *logical* truth. Further, each of its instances  $\exists_{16} y y \neq a$  will contain  $a$  inessentially, and so will qualify as a logical truth by Quine's definition, just as it qualifies as logically analytic under Bolzano's. To be sure, a philosopher of a very different persuasion (but probably not Quine!) might argue that these are no *contingent, empirical* truths, but are *necessary*. But that brings no respite, since it leaves untouched the central point, which is that they are surely not *logically* necessary or *logically* true—so that Quine's definition, and hence the Frege-Quine definition of analyticity which rests upon it, is in as bad a shape as Bolzano's.

a base class of analytic propositions, avoiding the Kneales' objection, by restricting the application of Bolzano's definition to statements free of singular terms. We might then, it seems, take care the remaining 'good' candidates, including analytic propositions whose expression involves the use of singular terms, by adding that a statement is analytically true if it is deducible from some statement(s) belonging to the base class. In short, the proposal is for two part definition:

- (1) A purely general statement is analytically true if it is true and contains at least one expression inessentially.
- (2) Any statement is analytically true if it is a logical consequence of some statement(s) analytically true by (1).

This proposal makes the status as analytically true of statements involving reference to particular objects derivative from that of analytic general statements, and so goes flat against our earlier proposal to secure the analyticity of statements like 'Whatever is yellow is coloured' (analyticities of the third kind) by taking a generalization to be analytic iff all its instances are. But since we have not shown that that is the only way to accommodate analyticities of the third kind, the present proposal remains, so far, a live option, and it is therefore worth considering whether, should it prove possible to accommodate analyticities of the third kind in some other way, it would be a viable option.

It would not be fair to object that the proposal is merely ad hoc. There is a well-established tradition of thought which has it that necessary truth has its source in relations among general concepts. The treatment of singular statements as analytic only when they are logically derivable from analytic general statements might be seen as a reflection of what is right in that admittedly somewhat sketchy thought.

It might also be objected that the proposal makes an unexplained use of the notion of logical consequence, and that when this is explained, the definition will turn out to be viciously circular. As against this, we may note that if this were a good objection, it would tell equally against the Frege-Quine definition. But in fact, it is unclear that an explanation of logical consequence need involve any appeal to the notion of analyticity. Standard explanations, to be sure, invoke the notion of necessary truth-preservation, or logical necessity, but neither is usually explained in terms of analyticity, and there is no compelling reason to think they must be. There is, however, a more serious objection.

Even if we can exclude counterexamples to the original definition by the emendation proposed, this does not dispose of the problem, because we can reduplicate the difficulty at the next level up. That is, just as we obtain counterexamples to Bolzano's original definition by exploiting synthetically true *first-level* generalizations to locate statements featuring singular terms that ought not to be, but are, counted as analytic by Bolzano's definition, so we can find synthetically true *second-level* generalizations whose truth ensures that uniform replacement of *first-level predicates* will not alter truth value—with the result that certain first-level generalizations that instantiate them rank as analytic under our revised definition, when they ought to come out as synthetic.

In fact, we can give an effective procedure for generating such higher-level counter-examples. We may assume that there are some merely synthetically true first-level generalizations. Let  $\forall xQx$  be any such. Then the first-level predicate  $Qx$  is true of every object. But then the second-level generalization  $\forall F\forall x(Fx \rightarrow Qx)$  is likewise synthetically true. Take any instance, say  $\forall x(Px \rightarrow Qx)$ . Then this will rank as analytic by clause (1)—for however we vary  $P$ , the resulting statement will be true, just because  $\forall F\forall x(Fx \rightarrow Qx)$  is.

## 5.2 *Necessitated inessentiality?*

It may be suggested that once we see *why* Bolzano's original definition is vulnerable to the kind of counter-examples we have discussed, it is not too difficult to see how his definition needs to be modified so as to exclude them. We can reformulate Bolzano's original definition in this way:

$S$  is analytic iff there is an expression  $u$  occurring in  $S$  such that where  $v$  is any other expression of the same grammatical type as  $u$ , the statement that results from  $S$  by substituting  $v$  for  $u$  throughout  $S$  is materially equivalent to  $S$

or more concisely:

$S$  is analytic iff  $\exists u(u \text{ occurs in } S \wedge \forall v(S[\frac{u}{v}] \leftrightarrow S))$

The present problem is that whenever  $S$  is an instance of some contingently true general statement, not only  $S$ , but also every other instance of that general statement will be true, as it happens, with the result that however

we vary the names or singular terms occurring in  $S$ , the resulting statements will always be alike in truth-value with  $S$ . So  $S$  will count as analytic. If, on the other hand,  $S$  is an instance of a contingently false general statement,  $S$  will not count as analytic, even if it happens to be true, because there will be some other instance of the general statement which is false, and so some singular term that can be substituted for a singular term occurring in  $S$  to yield a statement different in truth-value from  $S$ . Clearly, however, whether a statement is or is not analytic ought not to depend in this way on what merely *happens* to be the case. What determines whether or not  $S$  is analytic should be not whether substitutions of the kind in question *do as a matter of fact* lead to a change in truth-value, but whether or not they *could* do so. This suggests that we should strengthen Bolzano's definition in the following way:

$S$  is analytic iff  $\exists u(u \text{ occurs in } S \wedge \Box \forall v(S[\frac{u}{v}] \leftrightarrow S))$

This small adjustment clearly suffices to block unwanted candidates such as instances of true, but only contingently true, generalizations, since while the parent generalization's truth ensures that uniform substitution on singular terms will preserve truth-value, its contingency means that it *need not* do so. Of course, anyone who sympathizes with Quine's scepticism about the intelligibility of intensional idioms (such as the necessity operator) as opposed to supposedly purely extensional ones (such as truth and uniform substitution) will find this strengthening unacceptable. But we have already seen that the hope that Bolzano's approach would enable us to give an account of analytic truth in purely extensional terms is doomed to frustration. So we may set aside that objection here. There is, however, a much more serious problem. Consider the proposition:

If *this ring* is pure gold, *it* is entirely composed of a substance whose atomic number is 79.

We may substitute any singular term we wish for the italicized words and the resulting proposition not only will, but must, be true—assuming, as we certainly may, that the generalization of which we have taken an instance is not only true, but true as a matter of metaphysical necessity. But while the generalization, and so each of its instances, is metaphysically *necessary*, none of these propositions is *analytically* true. In short, the proposed emendation, as it stands, precipitates a collapse of the distinction between

metaphysical necessity and analyticity. The trouble lies with the unqualified or indiscriminate use of the necessity operator. In order to get the extensionally correct result, we would need somehow to specify that  $\Box$  is to express the right kind of necessity—one grounded purely in senses, or concepts—and it is quite unclear how we could do so without using the very notion we are trying to explain.

#### 6. *Over-extension (2)—the embedding problem*

We should now take note of a further serious problem with Bolzano's original definition, when it is viewed as an attempt to capture the traditional conception of analyticity as truth-in-virtue-of-meaning and as a potential improvement on the Frege-Quine definition. Let  $p$  be any proposition which qualifies as analytic by Bolzano's definition in virtue of containing the idea  $i$  inessentially, and consider its conjunction with  $q$ , where  $q$  is any intuitively synthetic proposition having the same truth-value as  $p$ , but not containing the idea  $i$  at all.<sup>14</sup> By hypothesis,  $p$  contains at least one idea which may be varied at will without yielding a proposition differing in truth-value from  $p$ . But then clearly the same must go for  $p \wedge q$ , given that  $q$  and  $p$  are alike in truth-value, and that  $i$  does not occur in  $q$ , so that varying  $i$  does not disturb  $q$ 's truth-value. Then  $p \wedge q$  will likewise qualify as analytic. Yet it clearly should not do so. It is true enough that if  $p$  is analytically false, so will be any conjunction of which it is a conjunct, so that Bolzano's definition gives the right result here. But suppose instead that  $p$  is true, and so analytically true according to the obvious way of defining analytic truth in Bolzano's terms. If  $q$  is true but synthetic, it seems clear that their conjunction should count as at best synthetically true. (cf. 'Married men are men and Handel outlived Bach').<sup>15</sup>

---

14. The point of the restriction is that if  $q$  contains  $i$ , then uniform replacement of  $i$  throughout  $p \wedge q$  by another idea  $i'$  may result in a proposition  $p \wedge q'$  which differs in truth-value from  $p \wedge q$ , just because  $i$  does not occur inessentially in  $q$ .

15. We should emphasize that we are not claiming that the problem is a problem for Bolzano, given his own purposes in giving his definition. That it is a problem for Bolzano is suggested by Jan Berg. Although Berg presents the problem as 'an objection from a modern viewpoint', conceding that 'Bolzano would probably not have considered [this] objection serious', he thinks it serious enough to add 'At any rate, this consequence of [his definition of broad analyticity] makes us concentrate our interest on the notion of logical analyticity' (see Berg (1962), 101, also his editorial introductions to Bolzano (1987), 18, and to Berg (1973), 18). We are grateful to Wolfgang Künne for the first two references. As Künne emphasizes (see Künne (2008a),

Essentially the same problem arises over disjunction. Let  $p$  be analytic, with  $i$  occurring inessentially, and let  $q$  be synthetic, materially equivalent to  $p$ ; and  $i$ -free. Then  $p \vee q$  will likewise be analytic. The problem, this time, arises when  $p$  is analytically false. Similar difficulties will arise with other embeddings of any proposition that is analytic by Bolzano's lights. In general terms, the problem is that if a statement  $A$  is analytic by Bolzano's definition, so will be any statement  $B$  which incorporates  $A$ , provided that the expressions in  $A$  in virtue of which it qualifies as analytic do not occur in  $B$  other than as parts of  $A$ , and  $A$  does not occur within a referentially opaque context in  $B$ .

It is easy to see that this embedding problem applies equally to each of our two attempts to rescue Bolzano's definition from the Kneales' objection. For if containing at least one idea inessentially is a sufficient condition for analyticity, as on the first proposal, then any conjunction one of whose conjuncts is analytic must be so also. And requiring, as on the second proposal, that uniform replacement of at least one idea should necessarily leave truth-value undisturbed equally clearly does nothing to alleviate the problem.

## 7. *Post mortem—and a better proposal?*

### 7.1 *Epistemologizing Bolzano*

Let's take stock. We have at this point two outstanding objections to Bolzano's proposal: the first, due to the Kneales, allows of a response only at the cost of the apparent circularity of invoking a notion of necessity in the explanans which itself promises to require explanation in terms of analyticity; the second—the embedding problem—seems to impose a disconnection between inessential occurrence and analyticity *prima facie* fatal to Bolzano's account.

Let us focus on the first of these difficulties. There is, as we have already emphasized (see page 6), a major difference between Bolzano's definition and Frege's: for Frege, the distinction between analytic and synthetic judgements, in line with that between a priori and a posteriori, is an *epistemological* one—the claim that a judgement is analytic is a claim about

---

248ff., esp. fn.48), while it is true that under Bolzano's definition, analytic propositions may be contingent and knowable only by empirical investigation, it is by no means clear that Bolzano would have found this consequence unwelcome or disturbing.

how it may be *justified*. By contrast, Bolzano defines the analyticity of a proposition simply in terms of the effect of varying some of its ingredient ideas upon its truth-value. But as the Kneales' objection brings out, inessential occurrence is no sure guide to analytic status, for it may have its source in some background contingencies. An obvious corollary is that the fact that an idea (or expression) occurs inessentially in a proposition (or statement) may itself be something recognizable a posteriori, via independent knowledge of the relevant contingencies. A further unwanted consequence is thus that Bolzano's definition threatens the traditional connection between analyticity and a priori knowability. In the light of all this, a natural and plausible response to the first of our problems is to 'epistemologize' Bolzano's definition: a proposition is logically analytic if it not only contains at least one idea inessentially and only logical ideas essentially, but is such that the fact that it does so can be recognized *simply* by relying upon one's grasp of those ideas or concepts involved in the proposition which cannot be varied freely; and, generalizing this, a proposition is analytic if it not only contains at least one idea inessentially, but is such that that fact can be recognized simply by relying upon one's grasp of those ideas or concepts involved in the proposition which cannot be varied freely.<sup>16</sup>

Accordingly, we may—as a first approximation—consider the following definition:

---

16. Interestingly, although his definition remains resolutely non-epistemological, Bolzano was sensitive to the kind of connection on which the proposed revision focuses. Commenting on his examples of logically analytic propositions (See page 4), Bolzano writes:

The examples of analytic propositions I have just cited are differentiated ... by the fact that nothing is necessary for judging the[ir] analytic nature ... besides logical knowledge, because the concepts that make up the invariant part of these propositions all belong to logic. (*Wissenschaftslehre* §148, Berg (1973), 193)

What is especially interesting here is Bolzano's saying that only logical knowledge is needed to recognize the *analytic nature* (rather than, as one might expect, the *truth*) of such propositions as those expressed by '*A* is *A*', '*An A* which is *B* is an *A*', etc. Recognizing the analytic nature consists, in his view, in seeing that certain ideas involved in the proposition can be varied in any way we please, and the result will be a proposition having the same truth-value as the original. It is this idea that the epistemologized version of Bolzano's definition we are about to consider takes up and generalizes. We are not, of course, suggesting that Bolzano himself harboured any thought that his original definition might be modified along these lines. On the contrary, he is firmly opposed to the introduction of any kind of epistemological considerations in defining analyticity. The non-epistemic character of Bolzano's definition is emphasized by Michael Dummett in Dummett (1991), 28ff.

*E-Bolzano 1* A statement  $A$  is analytic iff (i)  $A$  contains at least one expression which can be freely varied without change of truth-value (ii) that fact can be recognized by anyone who understands the remaining, non-variable expressions composing  $A$ , and grasps the semantic significance of its syntax.<sup>17</sup>

This modified definition avoids the Kneales' objection, and finesses the need to modalize in response to it. The objection exploits *empirical inessentiality*—the fact that an expression may indeed occur inessentially in a statement, but only courtesy of that statement's being an instance of some true empirical generalization. Where a statement does contain an expression which can be varied without change of truth value, but *only* because that statement is an instance of a true empirical generalization, grasp of the remaining expressions composing the statement precisely does *not* suffice to enable one to recognize that there is an expression which can be varied without disturbing truth-value. To know that, in such a case, one would need to know that the parent generalization is true, and mastery of the expressions involved in the candidate statement, though necessary, is insufficient for such knowledge.

However, while this modification escapes the Kneales' objection and preserves the principal potential advantage of Bolzano's original definition—of enabling us to see logical analyticity as a special case of a more general phenomenon, thereby avoiding the necessity of viewing the analyticity of logically true statements as a matter of direct stipulation, as on the Frege-Quine definition—it does nothing to alleviate the other major difficulty we found with Bolzano's original, viz. the embedding problem. Epistemologizing Bolzano's definition in the way indicated does not help. For if  $A$  is analytic in virtue of containing an expression  $e$  which can be varied freely without altering the truthvalue of  $A$ , and  $B$  is any longer statement incorporating  $A$  but containing no additional occurrences of  $e$ , anyone who understands  $B$  will be able to recognize that it contains  $A$  as a part, and contains no additional occurrences of  $e$ , and so will be able to recognize that  $B$  contains  $e$  inessentially.

---

17. Here and subsequently we treat analyticity as a property of *statements* (interpreted sentences), rather than *propositions*. When this is done, it is crucial to emphasize that the basis on which inessential occurrence is to be recognizable includes not only understanding of the statement's remaining, non-variable expressions, but also grasp of its syntax. We have included this last requirement here, but, in the interests of brevity, we will often leave it to be understood in the sequel.

## 7.2 *The embedding problem solved*

The embedding problem shows that even if Bolzano's definition leads us to count relatively simple statements as analytically true, or analytically false, just when they would be so classified in accordance with the traditional conception, it is liable to go badly astray when applied to more complex statements embedding them. Why is this? An obvious thought is that the problem reflects an important discrepancy between analyticity in Bolzano's sense and the traditional conception associated with the notions of truth/falsehood-in-virtue-of-meaning. The former is, as we might put it, *upwards-hereditary*, in the sense that the result of incorporating a Bolzano-analytic statement as part of a more complex statement must likewise be Bolzano-analytic, provided only that the remainder of the containing statement is free of further occurrences of the expressions occurring inessentially in the Bolzano-analytic part. But the traditional notion of analyticity clearly lacks this property. To take the simplest and most obvious examples, while the analytic falsehood of one conjunct suffices for that of the conjunction as a whole, a conjunction is analytically true only if both conjuncts are so; and while the analytic truth of one disjunct suffices for that of any disjunction incorporating it as a disjunct, a disjunction is analytically false only if both disjuncts are.

This initial diagnosis suggests that we might solve the problem by giving a recursive definition, using Bolzano-analyticity (or rather, our epistemologized version of it) to characterize a suitable base class, and using the recursive clauses to impose suitable requirements on the components of complex statements. Such a recursive definition can indeed be given, and in an appendix, we illustrate how this may be done for a first-order language. There is, however, another shortcoming which reflection on the embedding problem discloses, and this suggests a rather different remedy, making no essential play with recursion.

As previously observed (see p.5), Bolzano's definition of analytic propositions covers both analytically true and analytically false propositions, making no distinction between them. By contrast, such a distinction is central to the traditional conception, which explains analytic truth as truth-in-virtue-of-meaning and analytic falsehood as falsehood-in-virtue-of-meaning. Of course, one could define notions of analytic truth and falsehood in terms of Bolzano's notion of analyticity together with the notions of truth and falsehood, and one could define a general notion of analytic proposition in terms of the traditional notions of analytic truth

and analytic falsehood. But there remains a crucial difference. Starting from Bolzano's definition, we obtain:

*A is analytically true* iff *A* is analytic and *A* is true, and *A is analytically false* iff *A* is analytic and *A* is false.

Starting from the traditional notions, we obtain:

*A is analytic* iff *A* is analytically true or *A* is analytically false.

But the resultant notions of analytic truth and analytic falsehood under the first definition are plainly not equivalent to analytic truth and analytic falsehood as traditionally understood. Indeed, they are not even co-extensive, since

Haydn outlived Mozart and if Bartok and Kodaly were compatriots, Bartok and Kodaly were compatriots

counts as analytically true in Bolzano's sense, whereas it is clearly not so according to the traditional conception.

It is a consequence of precisely this divergence between Bolzano's notion and the traditional one that epistemologizing Bolzano's definition, as suggested in the preceding sub-section, does nothing to solve the embedding problem. Recognizing that a statement is analytic in the sense that it contains at least one expression which may be varied without change of truth-value is consistent with total ignorance of the statement's truth-value. In particular, someone who can recognize that the right conjunct in our last example is Bolzano-analytic is in position to recognize that the whole conjunction is so. Of course, this ability consists with total ignorance of the fact that the conjunction is, as it happens, true; but that is of no matter, since knowledge of truth-value is not required for knowledge of Bolzano-analyticity.

This suggests a quite different way of dealing with the embedding problem: emend E-Bolzano 1 to deal separately with analytic truth and analytic falsehood, and require recognition of truth-value as well as recognition of inessential occurrence. This yields

*E-Bolzano 2.1* A statement *A* is *analytically true* iff (i) *A* is true, (ii) *A* contains at least one expression which can be freely varied without change of truth-value, and (iii) that (i) and (ii) both

hold can be recognized by anyone who understands the remaining, non-variable expressions composing  $A$

*E-Bolzano 2.2* A statement  $A$  is *analytically false* iff (i)  $A$  is false, (ii)  $A$  contains at least one expression which can be freely varied without change of truth-value, and (iii) that (i) and (ii) both hold can be recognized by anyone who understands the remaining, non-variable expressions composing  $A$

Clearly, this adjustment is by itself enough to dispose of the embedding problem, without need for the complications of a recursive definition. Suppose, for example, that  $A$  is analytically true (i.e. meets conditions (i)–(iii) above), and consider its conjunction with any synthetic statement,  $B$ . Even if  $B$  is true, the fact that it is, and hence the fact that  $A \wedge B$  is true, will *not* be recognizable solely on the basis of competence with the remaining, essentially occurring, expressions in the conjunction. To be sure, should  $A$  be analytically *false*, then anyone competent in the use of the relevant expressions will be able to recognize that conditions (i)–(iii) are met with respect to its conjunction with any other statement, so that the conjunction will qualify as analytically false—but that is as it should be. Oppositely,  $A$ 's analytic truth will suffice for that of any disjunction  $A \vee B$ ; but rightly so; while should  $A$  be analytically false, this will not suffice to force analytic falsehood on its disjunction with arbitrary  $B$ ; for  $B$  may well be true. Other sentential compounds likewise raise no problem.

It remains to modify the foregoing proposal to accommodate analyticities of the third kind. These may be captured by modifying our definition in the way previously envisaged with Bolzano's own definition (See p.3). We propose:

*E-Bolzano 3.1* A statement  $A$  is *analytically true* iff (a) (i)  $A$  is true, (ii)  $A$  contains at least one expression which can be freely varied without change of truth-value, and (iii) that (i) and (ii) both hold can be recognized by anyone who understands the remaining, non-variable expressions composing  $A$ , or (b)  $A$  is a universal generalization whose instances are all analytically true or an existential generalization at least one of whose instances is so

with a similar adjustment for the definition of analytic falsehood.

## 8. *Analytic knowledge, epistemic and 'metaphysical' analyticity*

### 8.1 *An obvious complaint answered*

The foregoing proposal is, however, open to a very immediate complaint<sup>18</sup>: that the requirements that *A* be true (false) and that it should contain at least one expression inessentially, feature within it as *separate*, so far quite unconnected conditions. This was forced, given that—in contrast to Bolzano—we are distinguishing analytic truth and falsehood in our definitions. All the same, if any kind of account of the *nature* of analyticity is to be attempted, as opposed to a mere putative characterization of the extension of the notion, it is imperative to say more about how the two conditions are supposed to interact. As things stand, there is nothing to forestall the impression that two distinct acts of recognition are implicated in the recognition of analyticity—recognition of truth-value, on the one hand, and recognition of inessential occurrence, on the other. Worse, indeed, once the first component—recognition of truth-value—is written into the definitions in the fashion illustrated, does not the additional clause requiring inessential occurrence—the distinctive feature of Bolzano's original definition—become a mere curlicue?<sup>19</sup> What does the requirement that understanding the target sentence should enable recognition that one or more expressions occur inessentially within it add to the notion of epistemic analyticity, proposed by Paul Boghossian over the last couple of decades (see Boghossian (1996), Boghossian (1997), and Boghossian (2003)), according to which analytic sentences are those whose truth-value can be recognized purely on the basis of understanding them?

At first blush, it must be admitted, little or nothing of significance. But the appearance is arguably deceptive. The recognition that, say, 'If Bolzano is in Prague, then Bolzano is in Prague' is true, and the recognition that it contains the sentence 'Bolzano is in Prague' inessentially, are not two separate feats of recognition. Rather, someone who understands the conditional sentence can recognize that it expresses a truth (assuming, of course, that the proper names do not shift reference between antecedent and consequent) precisely because they know, in virtue of their understanding of the conditional construction, that provided the same sentence figures as antecedent and consequent, the conditional will be true, no matter what

---

18. The same complaint applies equally to the recursive definition described in the appendix, as the reader may easily verify.

19. In its English sense; apparently in mid-19th century American English it means a caper.

sentence that is. One recognizes truth, in such a case, by way of recognizing inessential occurrence. The point is no peculiarity of logically analytic truths. The same goes for any other minimally analytic truth. Consider, say, the proposition that if George is a brother, he is a sibling.

No one whose working vocabulary includes both ‘brother’ and ‘sibling’ needs to know anything about George in order to know that this proposition is true. Nor need they recognize that ‘If George is a brother, he is a sibling’ may be transformed into the logical truth, ‘If George is a male sibling, he is a sibling’. The ability to recognize the truth of the proposition can be entirely parallel to that of someone competent to recognize the truth of the proposition that if Mozart’s socks are yellow, they are coloured—where there can be no question of recognition proceeding through transformation into a logical truth, because there is no such transformation. One can recognize the truth of this particular proposition because one knows, simply in virtue of a competence with the terms, that no matter what term fills both gaps in the schematic sentence ‘if ... is a brother, ... is a sibling’, the resulting sentence will express a truth. And so, *mutatis mutandis*, for the proposition about Mozart’s socks. And of course, the same goes for recognition of minimal analytic falsehood. Anyone competent in the use of sentential negation and conjunction can recognize that ‘Cats are mammals and cats are not mammals’ must (assuming, of course, no relevant ambiguity) be false because they know, courtesy of their competence with the terms, that no matter what declarative sentence, true or false, occupies both gaps in the schema, ‘... and it is not the case that ...’, the result will be false.<sup>20</sup>

What these examples illustrate, we claim, is indeed the essentially *schematic* character of knowledge of analyticity—analytic truth or falsehood—in the basic (non-inferential) case. This is the insight that Bolzano’s definition—or at least his emphasis on inessential occurrence—contains. To recognize that the proper names occur inessentially in ‘If Haydn’s birth preceded Mozart’s, then Mozart’s followed Haydn’s’ is, in effect, to recognize that, no matter which terms fill the gaps in the schema ‘If \_\_\_’s birth preceded . . . ’s, then . . . ’s followed \_\_\_’s’, provided the same term fills the first and last, and the same term the second and third, the resulting statement will, and indeed must, have the same truth-value as our state-

---

20. No question is begged here against dialetheists. They do not deny that contradictions are always false. It is just that they think, and presumably claim to know, that some are true as well. They can avail themselves of this explanation of our recognition of their falsehood. That leaves the task of explaining putative knowledge of their truth. But that is none of our business.

ment about Haydn and Mozart. The same goes for recognition of inessential occurrence in analytic falsehoods, such as ‘Mozart’s last symphony was composed before Haydn’s, and Haydn’s last before Beethoven’s, but Beethoven’s last was composed before Mozart’s’. Recognition of analytic truth is, or centrally involves, recognition that a certain schema always yields a truth, on uniform insertion of suitable expressions in its gaps; and so, *mutatis mutandis*, for recognition of analytic falsehood.

Care is needed, though, in expressing the point. One does not—at least in general, if not invariably—*first* recognize that a sentence contains inessential occurrences of one or more expressions and *then*, purely on that basis, move to recognition that the results will be invariantly true (or invariantly false) no matter what expressions replace them. It is not that one recognizes analytic truth, or falsity, merely *by* recognizing inessential occurrence. To recognize that a sentence contains certain expressions inessentially need afford one no more than the knowledge that its truth-value will not depend on the semantic value of those expressions; exactly that was the gist of the embedding problem. The schematic character of analyticity is rather this: that in recognizing that a sentence is analytic, one recognizes that, such are the meanings of some (essentially occurring) expressions in it, and such is the semantic composition of the sentence as a whole, the sentence will—indeed, must—invariantly express a truth (or invariantly express a falsehood) no matter what the semantic values of the remaining (inessentially occurring) expressions it contains. In effect, the proposal is that the root of the notion of analyticity is a property not of truth-apt sentences in general but of *open* sentences: a property (the *Bolzano property*) which holds in virtue of the syntax and the semantic values of the expressions they contain and which ensures invariance of truth-value no matter whether we close them by instantiation or by universal generalization.<sup>21</sup> Thus neither ‘If Mozart’s stockings are yellow, they are coloured’ nor ‘Everything yellow is coloured’ is prior, in point of analyticity, to the other; rather each is posterior to the schema, ‘If ... is yellow, it is coloured’. The ground of our recognition of the analyticity of both the former, it is proposed, is the schematic knowledge expressed by the latter.<sup>22</sup>

---

21. The idea that for Bolzano, analyticity is a property of propositional forms is suggested in Lapointe (2011), see 62ff.; but this is not easily squared with what Bolzano himself says and conflicts with a more orthodox interpretation—see Künne (2008a), 233ff. We take no stand on this exegetical issue.

22. There is some delicacy with the point, since we are not saying, of course that no one can recognize analyticity whose language does not contain the resources for the expression of

We thus arrive at the following modification of *E-Bolzano 3*:

*E-Bolzano 4-Schematic* A schema  $S(\eta)$  is analytic iff (i) where  $S(e)$  results from uniform replacement of  $\eta$  throughout  $S(\eta)$  by any expression  $e$  of the type of  $\eta$ ,  $S(e)$  is always true, or always false, and (ii) that (i) holds is recognizable by anyone who understands  $S(\eta)$ .

*E-Bolzano 4.1* A statement  $A$  is *analytically true* iff (a)  $A$  is an instance of an analytic schema whose instances are always true or (b)  $A$  is a universal generalization whose instances are all analytically true or an existential generalization at least one of whose instances is so.

*E-Bolzano 4.2* A statement  $A$  is *analytically false* iff (a)  $A$  is an instance of an analytic schema whose instances are always false or (b)  $A$  is a universal generalization one of whose instances is analytically false or an existential generalization all of whose instances are so.

It would, we think, be unwise to claim that this modification takes care of all statements which might plausibly be reckoned analytically true or analytically false—in effect, that any analytically true statement is either an instance of an analytic schema, or is obtainable from such schemata by universal or existential generalization. Indeed, as far as English and other natural languages are concerned, it seems clear that this is not so. An interesting class of exceptions can be illustrated by examples such as:

Red is a colour  
Red is different from green  
Temporal precedence is a transitive relation

These examples exploit what we might call higher-order singular terms corresponding to first-level predicates—the nouns ‘red’ and ‘green’ corresponding

---

schemata. The claim is that what is recognized, when someone recognizes the analyticity of ‘If George is a brother, he is a sibling’, or ‘Anything yellow is coloured’, is something which, had they the appropriate expressive resources, could be formulated by means of a suitable claim about an open sentence. This should not seem uncomfortable unless one takes it that a subject’s knowledge is everywhere bounded by the resources they have for its expression.

to the predicates ‘... is red’ and ‘... is green’, and the abstract noun phrase ‘temporal precedence’ corresponding to the relational predicate ‘... temporally precedes\_\_’. How such examples are to be handled is a matter of some interest. It would distract us too much from our central line of argument to pursue this question here. We discuss it briefly in an appendix.

## 8.2 *Non-epistemic analyticity*

It is noteworthy that this proposal immediately provides resources sufficient to respond to Boghossian’s recently influential critique of what he termed the “metaphysical conception” of analyticity—the notion encapsulated in the idea of truth-purely-in-virtue-of-meaning. Boghossian complains that, taken at face value, the latter notion is incoherent: that no sentence can be true purely in virtue of its meaning. For any sentence *S*, if *S* is true, it will be because for some proposition *p*, *S* expresses *p* and it is a fact that *p*. A contribution from the world, or the facts, is always required even if the contribution is assured.<sup>23</sup> It is natural, however, to feel some discomfort with Boghossian’s own response to his point: the proposal to scrap the metaphysical notion altogether, in favour of an epistemic one, whereby a sentence ranks as epistemically analytic just in case an understanding of it provides a sufficient basis for recognition of its truth (or falsity).<sup>24</sup> For bracketing any scepticism whether that there are indeed such sentences, it could hardly be the last word about them to characterize them in that—purely epistemic—way. If grasp of a sentence’s meaning puts a subject in position to recognize its truth, there has to be something about its meaning in virtue of which that is so.<sup>25</sup> The proper conclusion is therefore only that, whatever that something is, it cannot be happily captioned as that the meaning of the sentence is such as to ensure its truth (falsity) with no contribution from the world.

*E-Bolzano-Schematic* now supplies a first-pass description of what the ‘something’ is: it is the property a sentence has when, such are its syntax

---

23. See, for example, Boghossian (1997), 335: “How could the mere fact that *S* means that *p* make it the case that *S* is true? Doesn’t it also have to be the case that *p*?”

24. cf. Boghossian (1997), 334: “On this [the epistemic] understanding, then, ‘analyticity’ is an overtly *epistemological* notion: a statement is ‘true by virtue of its meaning’ provided that grasp of its meaning alone suffices for justified belief in its truth”.

25. Acknowledging that analyticity cannot satisfactorily be conceived purely epistemically carries no commitment to any particular view about its source, much less a commitment to a realist or ‘metaphysical’ view. Hence our preference for the colourless term ‘non-epistemic analyticity’ over Boghossian’s more florid ‘metaphysical analyticity’.

and the meanings of the expressions essentially occurring in it, the open sentence resulting from the deletion of all inessentially occurring expressions and/or quantifiers is such as to generate a truth (falsehood) no matter how it is completed. This account finesses any threat of ‘marginalisation’ of the world in the process of the determination of the truth-values of analytic sentences, since analyticity is not now, in the first instance, a property of truth-apt sentences at all. We may of course extend the scope of the epithet, ‘analytic’, to encompass sentences resulting from analytic matrices by substitution or quantification into their argument places, by means of such further definitions as *E-Bolzano 4.1* and *E-Bolzano 4.2*. But then the truth-value of an analytic sentence is determined, just as it should be, *both* by meaning—the meaning of the open sentence from which it results, *and* by the world—in delivering the semantic values, necessary if it is to have a truth-value at all, of the particular inessentially occurring expressions it contains.<sup>26</sup>

### 8.3 *Concluding remarks*

So, for a theorist who wishes to salvage a metaphysical—better: non-epistemic—notion of analyticity, that may seem like progress. A caveat is immediately needed, however, since the characterization just offered over-extends to embrace, ‘If  $x$  is composed of water,  $x$  is composed of  $H_2O$ ’ as well as ‘If  $x$  is yellow,  $x$  is coloured’. And now it is tempting to think that the needed distinction can only be that, in the latter case, grasp of meaning supplies a complete basis for recognition of the invariance of truth-value of instances while in the former it does not. To exclude the unwanted cases, then, capturing just the traditionally analytic and excluding the necessary a posteriori, it seems that we must still characterize analyticity epistemically, as a property of the meanings of open sentences in virtue of which, unsupplemented by other information, it can be recognized ... etc. And this, it may well seem, still cannot be the last word; it cannot be that *all* there is to say about the property is that it sustains the relevant epistemic feat; there has to be an explanation of how it is sustained, of what it is about the matrices in question that enables one who understands them to recognize that their instances are invariant in truth value.

---

26. A rather different response to Boghossian’s *two-factor* argument, as she labels it, is advocated by Gillian Russell (in Russell (2008), 31–7). For a brief discussion of it, see Appendix 2.

Accordingly, a properly non-epistemic notion of analyticity must, it seems to us, find use for the notion of *grounding*: specifically, for different ways in which the possession by an open sentence of the Bolzano property may be underwritten. The invited distinction is very much along traditional lines: analytic sentences are instances, or generalizations, of matrices whose possession of the Bolzano-property is grounded purely in the senses of the expressions they contain, and in their *syntax*; other cases, like the water-H<sub>2</sub>O example, possess the Bolzano property in virtue of aspects of the essential nature of the semantic values of the expressions they contain essentially.

So in the end we arrive at a well-visited staging post on the road to vindication, or repudiation, of the notion of analyticity. Further progress from here, if possible at all, will require four things: consolidation of the notion of sense, explication of the notion of ground, an explanation of how the Bolzano-property can indeed be grounded in sense, and an explanation of how that fact can be non-inferentially recognized. Misgivings about any of these projects will continue to fuel scepticism about the notion. But the utterly convincing intellectual phenomenology of the usual stock of basic examples will continue to fuel resistance to that scepticism. We do not attribute to Bolzano any special insight into how the deadlock might be broken. But we do think that his ideas contain a contribution to the proper formulation of the problem that later discussion lost sight of. That is what we have tried to outline here.

### *Appendix 1: A recursive solution to the embedding problem*

As observed in 6.2, it is plausible to think that the embedding problem might be avoided by recasting our epistemologized version of Bolzano's definition as a recursive definition. Such a definition must, of course, assume a quite detailed analysis of the structure of the language to which it is to be applied, taking into account all the ways in which complex sentences may be constructed out of simpler ones. Although we know of no convincing reason to doubt that such an analysis may be given for natural languages such as English, we are certainly not in a position to provide one. We shall therefore address ourselves to a much more modest task—describing how a suitable recursive definition may be constructed for a schematic first-order language.

We assume, then, a first-order language comprising the usual truth-functional sentential operators together with universal and existential

quantifiers binding individual variables. The language will have a stock of first-level predicates of varying adicity, along with a stock of singular terms, from which the simplest sentences of the language may be formed.

Our first task in implementing this suggestion is to circumscribe a suitable base class of analytic statements. This is less straightforward than might be anticipated. We cannot take the base class to comprise just atomic or logically simple statements, since there are complex statements—e.g. the statement that if Mozart's socks are yellow, they are coloured—which we wish to count as analytic but which do not inherit their analyticity from that of their components. Indeed, it is far from obvious that there are *any* logically simple analytic statements. But if any complex statements are to be included in the base class, we must take especial care to block inclusion of any which simply re-introduce the embedding problem. This can be accomplished by taking the base class to comprise just those statements which, in addition to satisfying the epistemological condition previously proposed (see *E-Bolzano 1*), meet the further condition that they contain no proper part which does so. Clearly this will exclude such monsters as 'If  $p$  then  $p$  and grass is green', whilst admitting such as 'If Mozart's socks are yellow, they are coloured'.

However, there is a more serious snag. As we have observed, Bolzano sought to define what it is for a statement to be analytic, without differentiating between analytic truth and analytic falsehood; and *E-Bolzano 1* follows him in this regard. This poses no direct obstacle to devising a suitable clause for *negation*—clearly  $\neg A$  will be analytic iff  $A$  is. But with the binary connectives we are stymied. What clause should we adopt for conjunction, for example? We can't say that  $A \wedge B$  is analytic iff  $A$  and  $B$  both are—for analytic falsehood of either conjunct alone suffices for that of the conjunction, regardless of the status and truth-value of the other. But we can't say that  $A \wedge B$  is analytic iff *one of*  $A$  and  $B$  is—for if one conjunct is analytically *true*, the conjunction is surely so only if the other is so as well. It is easily verified that similar difficulties preclude any satisfactory clauses for the other connectives. The moral is clear. We must, after all, define analytic truth and falsehood separately.

Accordingly, we give a two part characterization of our base class. We abbreviate 'analytically true' and 'analytically false' to '*a*-true' and '*a*-false' respectively. We then define:

*A* is *minimally a-true* iff (i) *A* is true, (ii) *A* contains at least one expression inessentially, (iii) the fact that (i) and (ii) are met can be recognized

by anyone who understands  $A$  and grasps the semantic significance of its syntax, and (iv) no proper subformula of  $A$  meets (i),(ii) and (iii).  $A$  is *minimally  $a$ -false* iff (i)  $A$  is false, etc., [as for  $a$ -true, with 'false' replacing 'true'].

The full definition of analytic truth and analytic falsehood may then be given as follows:

- If  $A$  is minimally  $a$ -true,  $A$  is  $a$ -true
- If  $A$  is minimally  $a$ -false,  $A$  is  $a$ -false
- If  $A$  is  $\neg B$ ,  $A$  is  $a$ -true iff  $B$  is  $a$ -false, and  $a$ -false iff  $B$  is  $a$ -true
- If  $A$  is  $B \wedge C$ , then  $A$  is  $a$ -true iff  $B$  and  $C$  are, and  $a$ -false if  $B$  or  $C$  is
- If  $A$  is  $B \vee C$ , then  $A$  is  $a$ -true if  $B$  or  $C$  is, and  $a$ -false iff both are
- If  $A$  is  $B \rightarrow C$ , then  $A$  is  $a$ -true if  $\neg B$  or  $C$  is, and  $a$ -false iff  $B$  is  $a$ -true and  $C$   $a$ -false
- If  $A$  is  $B \leftrightarrow C$ , then  $A$  is  $a$ -true iff  $B \rightarrow C$  and  $C \rightarrow B$  are, and  $a$ -false one of them is  $a$ -true and the other  $a$ -false
- If  $A$  is  $\forall vB(v)$ , then  $A$  is  $a$ -true iff for every  $t$ ,  $B(\frac{v}{t})$  is, and  $a$ -false iff for some  $t$ ,  $B(\frac{v}{t})$  is
- If  $A$  is  $\exists vB(v)$ , then  $A$  is  $a$ -true iff for some  $t$ ,  $B(\frac{v}{t})$  is, and  $a$ -false iff for every  $t$ ,  $B(\frac{v}{t})$  is
- Otherwise  $A$  is neither  $a$ -true nor  $a$ -false

We have conditionals *only*, not biconditionals, in the clauses for  $a$ -truth for  $\vee$  and  $\rightarrow$ , and the clause for  $a$ -falsehood for  $\wedge$ , because statements with these operators as principal may qualify as minimally analytic; for example:  $Fa \vee \neg Fa$ ,  $Fa \rightarrow Fa$ ,  $Fa \wedge \neg Fa$ , as well as more interesting examples which are not logically analytic, such as ' $a$  is red  $\rightarrow a$  is coloured', etc. Instances of analyticities of the third kind qualify in precisely this way, while their parent general analyticities qualify by the clauses for the quantifiers.

There is no obvious obstacle to extending a definition along these lines to richer and expressively more powerful languages, involving higher-order quantification, or modal and perhaps other non-truthfunctional operators. However, we shall not pursue such extensions here.

*Appendix 2: Gillian Russell's response to the Two-Factor objection*

Boghossian asked (Boghossian (1997), 335): “How could the mere fact that  $S$  means that  $p$  make it the case that  $S$  is true? Doesn't it also have to be the case that  $p$ ?” The two-factor objection draws on the platitude that when a sentence is true, its being so is a function both of what the sentence means and how the world stands in relevant respects. According to Gillian Russell (Russell (2008), ch.1), one may coherently respond to it thus: to maintain that the truth-values of some sentences are fully determined by their meanings is not to be committed to denying that the world plays a part in determining their truth-value. Of course, that claim is incoherent, unless there are different kinds of determination, or senses of ‘determine’, in play. But so, she claims, there are.

To give her strategy some independent plausibility, Russell draws an analogy with multiplication. When one of the factors,  $a$  and  $b$ , is 0, their product  $a \times b$  is likewise 0, no matter what the value of the other factor. Supposing  $a = 0$ , we might say that the product is wholly determined by  $a$ . But this does not oblige us to deny that it results from multiplying *two* numbers—after all, without the other factor,  $b$ , there would be no product at all!

So it is, Russell argues, with the determination of a sentence's truth-value by the two factors of sentence-meaning and the state of the world. In general, neither factor wholly determines truth-value. But in the case of analytic sentences, just as with multiplication by 0, one factor—the sentence's meaning—by itself wholly determines the sentence's truth-value. However, this does not oblige us to deny that the other, worldly, factor plays a part, any more than we are obliged to deny that the other factor plays a part in multiplication by 0. It is just that, whatever the other factor is, we get the same result. Still, the result is the product of *two* factors, not one.

To underwrite this response, Russell offers some distinctions. Let  $f$  be an  $n$ -ary function, and let  $x_i \dots x_k$  be some or all of the  $n$ -tuple of arguments  $x_1 \dots x_n$ . Then, first,  $x_i \dots x_k$  *fully determine* the value  $y = f(x_1 \dots x_n)$  if, for any  $n$ -tuple of arguments  $x'_1 \dots x'_n$  which coincide with  $x_1 \dots x_n$  over  $x_i \dots x_k$ ,  $f$  has  $y$  as value, regardless of the remaining arguments, if any. Second, an *argument-place*  $i$  in the sequence of argument-places  $\langle 1 \dots n \rangle$  *partially determines* the *value-place* of the function, if there are sequences of arguments  $x_1 \dots x_n$  and  $x'_1 \dots x'_n$  which differ in exactly their  $i^{\text{th}}$  place, such that  $f(x_1 \dots x_i \dots x_n) \neq f(x_1 \dots x'_i \dots x_n)$ . Finally, an argument  $x_i$  *redundantly determines* the

value,  $y$ , of the function if  $f$ 's  $i^{\text{th}}$  argument-place partially determines its value-place, but there is no sequence of arguments  $x'_1 \dots x'_n$  which differs from  $x_1 \dots x_n$  in and only in the  $i^{\text{th}}$  place, and delivers a value  $y' \neq y$ . These definitions ensure, as of course they are precisely designed to do, that some of the arguments to a function may fully determine its value while another of its arguments redundantly determines that value.<sup>27</sup>

Applying these distinctions to the case in hand, Russell's proposal is that just as the binary function, multiplication, maps pairs of numbers to numbers, so there is a binary function—she labels it  $\mathcal{M}$ —which maps pairs whose first member is a sentence-meaning (or proposition) and whose second is a state of the world to truth-values (see Russell (2008), 35). In the case of an analytic sentence  $S$ , the truth-value is *fully* determined by the first factor,  $S$ 's meaning—that is,  $\mathcal{M}$ 's value for the pair  $\langle m, w \rangle$ , where  $m$  is  $S$ 's meaning, will be the truth-value True, no matter what the value of  $w$ , the state of the world, may be. But this does not mean that the state of the world plays no part, for it does 'redundantly' determine the truth-value. Thus, she claims, the two-factor objection can be answered.

As the discussion in the main body of this paper will have made clear, we are sympathetic to the spirit of Russell's proposal to preserve a non-epistemic notion of analyticity from Boghossian's objection. But we are doubtful about the specific tack she takes. The most immediate complaint is that it fails properly to address the central point of the two factor objection, viz. the claim that the world not merely invariably plays *some* part in determining truth-value, but that when any sentence  $S$  is true, what makes it so is the fact that for some  $p$ ,  $S$  says that  $p$ , and—specifically—that it is the case that  $p$ . That is, it is—on the worldly side—not just any old fact, but the particular fact that  $p$  which combines with  $S$ 's meaning to deliver the truth-value. In the case of a multiplication for the form,  $0 \times a$ , the  $a$ -argument is there simply to make up the numbers, as it were—since, as remarked, without it or something in its place, there would be no output value. But in the case of the determination of a sentence's truth-value by its meaning and the 'state of the world', the latter has to have a specific character—specifically, as demanded by the meaning of  $S$ , it has to incorporate the fact that  $p$ —if the value, True, is to result. So Russell's analogy breaks down: the suggestion that states of the world and the non-zero fac-

---

27. Note that while full and redundant determination are relations on *arguments* to functions and their values, partial determination relates argument- and value-*places*. If partial determination were defined on arguments instead of argument-places, the definition of redundant determination would be flatly inconsistent.

tors in multiplications by zero are alike in ‘redundantly determining’ the truth-values of analytic sentences and the product of the multiplications respectively, masks a crucial difference in the roles they play.

To sharpen the complaint, let’s take the worldly argument of the  $\mathcal{M}$ -function to be a specific state of affairs—a state of the world in the sense in which, as we write, that Obama is the US President, or that there is currently an intensification of hostilities in Gaza, are states of the world. Let’s write the curly-bracketed  $\{Vivaldi\ died\ in\ Venice\}$  to denote the state of the world consisting in Vivaldi’s dying in Venice, and let  $\{S\}$  in general denote both actually obtaining state of affairs and merely possible ones (as in our example, Dear Reader—Vivaldi actually died in Vienna). Let the square-bracketed  $[Vivaldi\ died\ in\ Venice]$  denote the meaning of “Vivaldi died in Venice”; and let’s write  $\mathcal{M}([S],\{S\}) = v$  to express that Russell’s  $\mathcal{M}$ -function has value  $v$  for a certain sentence-meaning and state of affairs as its arguments. Clearly, we should have  $\mathcal{M}([S],\{S\}) = \text{True}$  and  $\mathcal{M}([S],\{\neg S\}) = \text{False}$ . But the crucial question is: What should be the value of our function for an arbitrary pair  $\langle [S],\{T\} \rangle$  where  $S$  and  $T$  are different sentences. What, for example, is  $\mathcal{M}([Vixens\ are\ female],\{Haydn\ invented\ the\ string\ quartet\})$ ? The question poses a dilemma. Either  $\mathcal{M}$  is defined for such ill-matched argument pairs, or it is not. Since Haydn’s invention of the string quartet has no bearing upon the sex of vixens, it is natural to declare  $\mathcal{M}$  undefined in this case. But if we say this here, we must say it everywhere except when first and second arguments are specified using the same sentence or its negation. And that would be hopeless for Russell’s purpose, since the only argument pairs for which an *analytic* sentence  $A$  will be defined will be  $\langle [A],\{A\} \rangle$  and  $\langle [A],\{\neg A\} \rangle$ —or perhaps, slightly less restrictively, for pairs in which the sentence used to specify the second argument expresses the same proposition as that used to specify the first. The result is that the desired contrast with non-analytic sentences is lost—since for exactly the same reason, we should deny that  $\mathcal{M}$  is defined for all other argument pairs, such as  $\langle [Vivaldi\ died\ in\ Venice],\{Haydn\ was\ born\ in\ Rohrau\} \rangle$  in which the state of affairs figuring as the second argument has no bearing on the truth-value of the proposition which figures as the first.

If, on the other horn, we insist that  $\mathcal{M}$  is defined for such pairs, and claim that analytic statements are precisely distinguished by the fact that when a statement  $A$  is analytically true,  $\mathcal{M}([A],\{B\})$  has the value True regardless of the state of affairs serving as its second argument, we thereby surrender all grip on the idea, encapsulated in the truth-meaning platitude underpinning the two-factor objection, that even in the case of an analytic

statement  $A$ , there is a *particular* worldly factor that is distinctively relevant to  $A$ 's truth, viz. the fact that  $A$ .

It is true that Russell herself seems to have in mind that the worldly factors which redundantly determine  $\mathcal{M}$ 's value in the case of analytic sentences are not states of the world understood in this particularistic way but are something more like possible worlds in the usual sense, i.e. *complete* ways for the world to be.<sup>28</sup> But this makes no difference. The objection does not go away. Rather, it resurfaces as the requirement that in order for a state of the world, globally so interpreted, to redundantly determine  $A$ 's truth-value as True, it has to be the case that any possible state of the world—any possible argument for the second place in the M-function—will incorporate the fact that  $A$ . This requirement has no counterpart in Russell's prototype of multiplication by zero—there is, as it were, no *zero-specific* requirement placed on the second factor in the multiplication if zero is to be the product, in the way that there is an  $A$ -specific requirement placed on the global worldly states if True is always to be the value. Moreover there now seems to be an imminent danger that, so far from making space for something like the traditional idea of truth-purely-in-virtue-of-meaning, the resulting proposal has the order of determination the wrong way round. It is down to the nature of zero that the result of multiplying it by any number is zero. Correspondingly, for a defender of the traditional idea, it ought to be down to the nature—the meaning—of an analytic truth  $A$  that the value  $\mathcal{M}$  gives for the pair consisting of it and any possible state of the world is True. But on Russell's proposal, with states of the world now globally understood, matters seem to run the other—wrong—way around: that is, it is only because every possible state of the word incorporates the fact that  $A$  that  $\mathcal{M}([A], \{...\})$  yields True whatever replaces the dots.<sup>29</sup>

28. Her term 'state of the world' for the second argument could, of course, be interpreted either way. But in giving particular examples (see Russell (2008), 35, fn.4), she uses  $w_\alpha$  and  $w_\beta$  explaining that ' $w_\alpha$  denotes the actual world, and  $w_\beta$  denotes a possible world in which snow is black'.

29. And of course it does not matter, for all that has so far been said, *why* every possible state of the world incorporates the fact that  $A$ —why there are no non- $A$  worlds. This is in effect why, in his review of Russell's book, Boghossian was able justly to complain that her proposed solution "still leaves us with the problem of distinguishing the merely necessary from the analytic, because we can equally well say that in the case of a necessarily true sentence, the truth-value is 'fully determined' by the meaning factor alone" (see Boghossian (2011), 371).

Russell is, of course, fully alive to the danger that her attempt to rescue the metaphysical notion of analyticity will lead to all necessary truths counting as analytic, and devotes a large chunk of her book (chapters 2 and 3) to addressing it. We have no space pursue that question

*Appendix 3: Higher-order singular analyticities*

As noted in the main text (see p.21), there are candidate analyticities, deploying what we called higher-order singular terms, which are not—at least not obviously—obtainable from analytic open sentences by instantiation or quantification, and which, therefore, pose a challenge to the essentially schematic conception of analyticity we have presented as a development of Bolzano’s central idea. As illustrative examples, consider:

- Red is a colour
- Red is a property
- Red is different from green
- Crimson is a determinate of red
- Temporal precedence is a relation
- Temporal precedence is transitive
- Addition is a function from numbers to numbers
- The natural numbers are closed under addition

These are all, intuitively, as good a range of candidates for epistemic analyticity as are any sentences. They are all, that is, such that it is tempting to say that someone who fully understands them is thereby put in position to know that they express truths. Yet how might they be accommodated by the schematic conception?

A comprehensive treatment of such examples is beyond the scope of the present discussion.<sup>30</sup> Here we merely outline what in our view (from

---

here, beyond observing that the point we have pressed does not depend upon its resolution. As far as we can see, the complications and revisions she introduces in these chapters are entirely driven by the need to avert the threatened collapse of analyticity into necessity, and have no bearing on the two-factor objection, which she appears to take to have been adequately answered in the preceding chapter. The ensuing complications do not materially affect that answer. True enough, the binary function  $\mathcal{M}$  from meaning-world pairs to truth-values is replaced (see Russell (2008), 53–7) by a quaternary function  $\mathcal{M}'$  whose arguments are a context of introduction, a context of utterance and a context of evaluation for an expression, along with what she calls a ‘reference determiner’ associated with it. In case the expression is a sentence, the  $\mathcal{M}'$ -function takes quadruples of such arguments to a truth-value. Crucially, however, the third argument—the context of evaluation—will still be, in effect, just as with the simpler  $\mathcal{M}$ -function, a ‘state of the world’. And that is enough to set up the complaint of the Appendix.

30. The use of higher-order singular terms seems to us to merit systematic study. It is, arguably, no mere an isolated curiosity; on the contrary, a case can be made that the introduction of such singular terms corresponding to predicates, relational, and functional expressions, often by more or less explicit kinds of nominalization, plays an indispensable rôle in semantic and

the standpoint of a non-sceptic about the notion of analyticity) should be made of them.

The natural first thought, if such cases are to be brought within range of the schematic approach, is to try to translate them into sentences whose analyticity is straightforwardly treatable in terms of that approach. The use of abstract nouns to express, in compressed form, what may more compendiously be expressed as a first-level generalization about concrete entities is common enough in natural languages. Thus we can say ‘Wisdom is a virtue’ when we might have expressed ourselves less concisely by saying ‘Anyone who is wise is, to that extent, virtuous’, or some such. No doubt many of the kind of examples illustrated might be brought within reach of the schematic conception by paraphrasing them as generalizations in this kind of way, which would then be covered by *E-Bolzano 4.1* or *E-Bolzano 4.2*. For instance, it might be proposed that our first and sixth examples above might be paraphrased as:

Anything which is red is, as such, coloured

and

If one event precedes another, and the second precedes a third, then the first precedes the third

We doubt, however, that the reductive paraphrase strategy can give an adequate account of all the problematic higher-order cases. Even if some of the examples may seem to admit of fairly natural and plausible transformation into lower-level generalizations, it is very doubtful that all do so—how, for example, should we paraphrase ‘Red is a property’, or ‘Addition is a function from numbers to numbers’. But more importantly, what exactly is paraphrase supposed to accomplish? The initial explanandum, remember, is the apparent *epistemic* analyticity of the kind of example illustrated. So is the idea that the *epistemic route* by which the (putative) analyticity of such examples is recognized *goes through* the suggested paraphrase? That seems to us quite implausible, even in the case of the examples like ‘Red is a colour’, where, on the contrary, it seems that once a speaker has acquired the use of the nouns ‘red’ and ‘colour’, she can directly recognize

---

ontological theorizing. But it raises some hard problems. For a discussion of some of them, see Hale and Linnebo (forthcoming).

the truth of ‘Red is a colour’, without the need for any detour through the proposed paraphrase.

The right approach, we believe, is to take the singularity of the examples seriously. Rather than essay to see them as some kind of idiomatic variants on lower-level generalisations, we should attempt to account for their distinctive epistemic status in a way that connects directly with their overt syntactic structure. ‘Red is a colour’ says something about the kind of property that the particular property, red, is; ‘Temporal precedence is transitive’ likewise says something about the character of the relation of temporal precedence, and so on. These are higher-order singular necessary truths, and a satisfactory account of their distinctive epistemic status should acknowledge them as such.

Singular necessary truths in general are not a rare bird. Since *Naming and Necessity*, philosophers have been very mindful of the kind of singular necessities typified by ‘Water is H<sub>2</sub>O’, ‘Heat is molecular motion’, ‘Saul Kripke is the son of Dorothy and Myer’, and so on. These are necessities of *essence*, broadly construed: propositions whose truth is determined by the essential nature of the referents of their subject terms. We suggest it is no different with the examples in the list above. ‘Red is a colour’ is true in virtue of the essential nature of the property, red; ‘Temporal precedence is transitive’ is true in virtue of the essential nature of the relation of temporal precedence. Yet the Kripkean examples are of necessities known *a posteriori*, while the higher-order singular statements on our list present, rather, as knowable *a priori*—hence their appearance as, intuitively, epistemically analytic. How then can there be any close comparison between the two kinds of case?

Well, the comparison can be saved while acknowledging the epistemic contrast, provided that the relevant essences, in the case of the higher-order singular statements, and unlike the Kripkean cases, are themselves knowable *a priori*: more specifically, provided that those aspects of the nature of the associated properties in virtue of which one who understands statements of the relevant kind incorporating higher-order singular terms is empowered to recognise them as true, are given with an understanding of the relevant singular terms. An account pursuing that thought may draw on what is standardly called an *abundant* conception of properties and relations.<sup>31</sup> According to the abundant conception, it suffices

---

31. The terminology of abundant and sparse properties originates in Lewis (1986). The general distinction is in Armstrong (1979). See also Bealer (1982), Swyer (1996), and Shaffer (2004). For a useful overview, see Mellor and Oliver (1997). Our own discussions include Hale and Wright (2009)], 197ff., 207ff., Hale (2013), §§1.12, 7.2, and Hale (2010), §§3.9.

for a predicate to express, or stand for, a property or relation that it be associated with a well-determined satisfaction condition, and the nature of the property or relation involved is, moreover, fully manifest in that satisfaction-condition. Thus the essence or nature of an abundant property or abundant relation does not lie beneath the surface, as it were, awaiting discovery by painstaking scientific investigation, but is open to view, and directly available to anyone who grasps the relevant satisfaction condition (see Hale (2013), §11.2).

In such cases, there is an a priori route to knowledge of essence, and this extends to support knowledge a priori of the higher-order analyticities illustrated by our list. A priori recognition that red is a colour rests, not on registering a relation of inclusion between the senses of the nouns 'red' and 'colour', still less on appreciating that it is equivalent to some analytic generalisation featuring the predicates, 'red' and 'coloured', but on knowing, as competence in its use requires, that the noun 'red' stands for the property attributed by the use of the corresponding adjective—the property whose essence is fully encapsulated in the satisfaction condition associated with the predicate '... is red', along with knowledge of a parallel fact about what the noun 'colour' stands for.

If this suggestion is broadly correct, we should recognise that we have to deal with two quite different kinds of analyticity. There is the phenomenon of analytic schematic generality, discerned by Bolzano, on which the main body of our discussion has concentrated; and there is a phenomenon of abundant property essence, illustrated by the present examples. Still, in view of the semantic relationship between (*n*-place) predicates and the corresponding higherorder terms—which of course needs a proper account—and supposing that an abundant conception of the referents of the latter, and the transparency that enjoins, is accepted, it remains that both phenomena are rooted *au fond* in the satisfaction-conditions of predicates, and thus in meaning.

Which is as it ought to be.<sup>32</sup>

---

32. We are grateful to Beau Mount and Filippo Ferrari for pressing the need for a treatment of the higher-order singular cases in discussion of these issues at the 2014 Northern Institute of Philosophy Summer School on the Foundations of Logic and Mathematics. A similar point was urged by Manuel Garcia-Carpintero in discussion at a workshop on analyticity in Lisbon. We are grateful, too, to Mark Textor for very helpful advice on the interpretation of Bolzano's work, to Keith Hossack for useful discussion of an earlier version, and to Jared Warren, and especially to Wolfgang Künne, for detailed written comments on our penultimate draft.

## REFERENCES

- Armstrong, David 1979: *A Theory of Universals*. Cambridge: Cambridge University Press.
- Bealer, George 1982: *Quality and Concept*. Oxford: Oxford University Press.
- Berg, Jan 1962: *Bolzano's Logic*. Stockholm: Almqvist and Wiksell.
- Berg, Jan (ed.) 1973: *Theory of Science*. Dordrecht: D. Reidel (part translation of Bernard Bolzano 1837).
- Boghossian, Paul 1996: "Analyticity Reconsidered". *Noûs* 30, 360–91.
- 1997: "Analyticity". In: Bob Hale & Crispin Wright (eds.), *A Companion to the Philosophy of Language*. Oxford: Basil Blackwell, 331–68.
- 2003: "Epistemic Analyticity: A Defense". *Grazer Philosophisches Studien* 66, 15–33.
- 2011: "Truth in Virtue of Meaning". *Australasian Journal of Philosophy* 89, 370–74. (Review of Gillian Russell 2008).
- Bolzano, Bernard 1837: *Wissenschaftslehre*. Sulzbach: J.E.von Seidel.
- 1987: *Bernard Bolzano Gesamtausgabe*, volume 12/1 of 1. Stuttgart/Bad Cannstatt: Frommann-Holzboog.
- Dummett, Michael 1991: *Frege Philosophy of Mathematics*. London: Duckworth.
- Frege, Gottlob 1974: *The Foundations of Arithmetic: A Logic-Mathematical Enquiry into the Concept of Number*. Oxford: Basil Blackwell. (English translation by J. L. Austin of Gottlob Frege 1884).
- 1884: *Die Grundlagen der Arithmetik. Eine logisch mathematische Untersuchung über den Begriff der Zahl*. Breslau: Wilhelm Koebner.
- George, Rolf (ed.) 1972: *Theory of Science*. Oxford: Basil Blackwell. (Abridged translation of Bernard Bolzano, 1837).
- Hale, Bob 2010: "The Bearable Lightness of Being". *Axiomathes* 20, 399–422.
- 2013: *Necessary Beings: An Essay on Ontology, Modality, and the Relations between them*. Oxford: Oxford University Press.
- Hale, Bob & Linnebo, Øystein (forthcoming): "Fregean Categories and the Problem of Expressibility".
- Hale, Bob & Wright, Crispin 2009: "The Metaontology of Abstraction." In: David Chalmers et al (eds.), *Metametaphysics*. Oxford: Clarendon Press, 178–212.
- Kneale, William C. & Kneale, Martha 1962: *The Development of Logic*. Oxford: Clarendon Press.
- Künne, Wolfgang 2008: "Analyticity and Logical Truth: From Bolzano to Quine." In: Wolfgang Künne (ed.), *Versuche über Bolzano/Essays on Bolzano*. Academia, Sankt Augustin, 233–304.
- Lapointe, Sandra 2011: *Bolzano's Theoretical Philosophy, History of Analytic Philosophy*. Oxford: Palgrave Macmillan.

- Lewis, David 1986: *On the Plurality of Worlds*. Oxford: Blackwell.
- Mellor, D.H. & Oliver, Alex 1997: "Introduction". In: D.H. Mellor & Alex Oliver (eds.), *Properties*. Oxford: Oxford University Press.
- Quine, Willard Van Orman 1953: "Two Dogmas of Empiricism" In Willard V. O. Quine (ed.), *From a Logical Point of View*. Harvard: Harvard University Press.
- 1966: "Carnap and Logical Truth" In: Willard V. O. Quine (ed.), *The Ways of Paradox*. New York: Random House, 100–125.
- Russell, Gillian 2008: *Truth in Virtue of Meaning*. Oxford: Oxford University Press.
- Shaffer, Jonathan 2004: "Two Conceptions of Sparse Properties". *Pacific Philosophical Quarterly* 85, 92–102.
- Strawson, Peter F. 1957: "Propositions, Concepts, and Logical Truths". *The Philosophical Quarterly* 7, 15–25.
- Swayer, Chris 1996: "Theories of Properties: from Plenitude to paucity". *Philosophical Perspectives* 10, 243–64.

ON HAVING A PROPERTY.  
CORRIGENDA IN BOLZANO'S *WISSENSCHAFTSLEHRE*.<sup>1</sup>

Wolfgang KÜNNE  
University of Hamburg

*Summary*

The hard core of Bolzano's conception of truth can be captured by the following biconditional: a non-compound non-quantified one-place predication expresses a true proposition (*Satz an sich*) if, and only if, the object that is referred to in that predication (really) has the property that is predicated of it. Thus the proposition that is expressed by the sentences 'Socrates is courageous', '*Socrate est courageux*', '*Sokrates ist mutig*' and many others is true just in case Socrates has courage. In this paper I shall focus on the sense of two words that occur not only in my characterization of the hard core of Bolzano's conception of truth but also in his own analysis (*Erklärung*) of this notion; namely, the sense of 'has (*hat*)' and the sense of 'property (*Beschaffenheit*)'.

1.

Let us begin with the three-letter word. In his logic Bolzano makes extensive use of the schemata 'A has (small-) b' and 'A is (capital-) B'. Whenever

---

1. *When I met Peter Simons for the first time (as one of three British philosophers who shared an infectious enthusiasm for what they called Austrian philosophy), he was a Lecturer at a place called the Bolton Institute of Technology, so this was at a very early stage of his academic career. After our first meeting in England, I received numerous beautiful handwritten letters on numbers and manifolds; he apparently did not mind my typewritten responses, and very soon we became friends. Already in those years Peter's work had an impact on mine, and I have profited from his writings ever since. We met very often when he lived in his home country, in Austria, and I am very glad that we have remained in close contact even after he went into exile. I seem to remember that in the early years of our acquaintance the German Brentano was the earliest of the Austrian philosophers whose work Kevin Mulligan, Barry Smith and Peter had studied. As soon as Peter had arrived in Edgar Morscher's Salzburg, this was bound to change, of course, and by now Peter is ready to confess: 'It has become my unshakeable opinion that Bolzano is the finest philosopher of the nineteenth century, bar none' (Simons 2014, 166). Hence I am certain that he will not be displeased when he finds in his amply deserved Festschrift a critical essay on some aspects of the great Bohemian's logic and metaphysics.*

its instances stand in systematic correspondence he uses them in the following way: the term in the ‘b’ position of the ‘has’ sentence is the nominalization of term in the ‘B’ position of the corresponding ‘is’ sentence. Thus the schemata turn into pairs like ‘Nathan has wisdom’ and ‘Nathan is wise’, or

- ( $\Sigma$ ) Socrates has courage
- (S) Socrates is courageous.

Bolzano takes it to be a matter of course that the partners in such couples express the same proposition (1837, II, 10;<sup>2</sup> 1841, 48). Ramsey agrees (1925, 60, 71), and so does Strawson who uses the same example: ‘The syntactical variation between “Socrates has courage” and “Socrates is courageous” is not more than that—a syntactical variation’ (1990, 318). (In spite of this prominent support I think that Bolzano’s propositional-identity claim is questionable, but I shall not question it here.<sup>3</sup>)

The proposition expressed by (S) is ‘more clearly (*deutlicher*) and more correctly (*richtiger*) expressed’ by ( $\Sigma$ ), or so Bolzano maintains (1837, II, 11). Why does he take ( $\Sigma$ ) to surpass (S) as regards clarity? I think his reason is similar to the reason logicians have for preferring ‘Plato is a philosopher, and Aristotle is a philosopher’ to ‘Plato and Aristotle are philosophers’. According to the pivotal Bolzanian truth-equivalence, (S) expresses a truth just in case the object referred to in (S) really has the property (S) ascribes to it. If that is the best way to specify the truth conditions of (S) then the greater clarity of ( $\Sigma$ ) consists in the fact that it contains a term that answers the question, ‘And which property is that?’. It is not obvious that this really *is* the best way of specifying the truth conditions of sentences like (S), but that is a question I shall not pursue here.

What about the other compliment Bolzano makes to ( $\Sigma$ ),—what is its greater (expressive) ‘correctness’ supposed to consist in? Bolzano believes that in sentences like (S) ‘the verb ‘to be’ is not used in its proper meaning (*eigentliche Bedeutung*), but only in that improper [i.e. figurative] meaning that it has as copula’.<sup>4</sup> It has its proper meaning, he claims, in sentences like

2. Quotations from *Wissenschaftslehre* (= 1837) and from Bolzano’s *Lehrbuch der Religionswissenschaft* (= 1834) refer to volume number and page number of the first editions.

3. For discussion see Künne 2006, 266–72 (‘Elementary Predications and their Quasi-Platonic Counterparts’).

4. 1837, I, 201. Cp. 1837, II, 19; 1935, 29f; 1841, 49. On the opposition ‘*eigentliche / uneigentliche Bedeutung*’ cp. 1837, III, 71.

‘God is’ and ‘I am’ where it is used to ascribe actuality (*Wirklichkeit, Seyn, Daseyn, Existenz*) to an object. (I shall return to the non-modal notion that these four words, interchangeably used by Bolzano, express in his mouth.) Suppose he is right. Then sentence ( $\Sigma$ ) is the ‘more correct’ formulation of the proposition that is also expressed by (S), since only ( $\Sigma$ ) is true under a *literal* reading: it employs no word in a figurative meaning. But it is not clear that the alleged figurativeness of the use of ‘is’ in (S) would justify a preference for ( $\Sigma$ ). Isn’t ( $\Sigma$ ) tantamount to ‘Socrates possesses courage’? And doesn’t that show that a juridical metaphor is in play?<sup>5</sup>

Is the use of ‘is’ as copula really figurative? One may very well wonder why the copulative use of this verb was, and is, vastly more common than its absolute use if the verb carries its proper meaning only in the absolute use.<sup>6</sup> At this point a look into Adelung’s dictionary of 18th century German is instructive. He illustrates the absolute use by a couple of excellent examples. (It is because of their linguistic peculiarities that I refrain from translating the passage.<sup>7</sup>) ‘*Seyn*’, he writes, is sometimes used in the sense of

Vorhanden seyn, wirklich seyn. *War je ein Wunsch, den mein Auge verrieth, den du nicht erfülltest*, Geßn[er]. Besonders absolute, die reelle Existenz, das Daseyn eines Dinges außer den Gedanken zu bezeichnen. *Es ist ein Gott. Gott ist. Gott ist von Ewigkeit her gewesen. Ich denke, darum bin ich. Daß ich jetzt bin, ist unverdiente Wohlthat des Schöpfers*, Gell[ert]. *Das Verlangen nach Glück verläßt uns nur in dem Augenblicke, da wir aufhören zu seyn.*

Er lebet, wie gar viel schließt dieses Wort nicht ein!

*Ihr Weisen, saget mir, heißt leben mehr als seyn?*

Hag[edorn].

In der dichterischen Schreibart bedeutet es oft nur[:] in der Reihe der sichtbaren Körper, der lebendigen Dinge, vorhanden seyn, leben. *Sie werden mich*

5. Bolzano himself occasionally switches from ‘A has b’ to ‘A possesses b’: see, for example, 1837, II, 26. In this respect Simons’ nickname for Bolzano’s ‘has’ is quite appropriate: ‘*Possessiv-Kopula*’ (Simons 1999, 15). At this point it may also be worth remembering the use of both ‘has’ and ‘property’ in contexts like ‘He has a small property in Kent’. In German there is no such use of ‘*Beschaffenheit*’, but once upon a time ‘*Eigenschafft*’ was tantamount to the juridical term ‘*Eigentum*’ (possession) rather than to the ontological term ‘quality’ (cp. Adelung s.v. ‘*Eigenschafft*’). An echo of the old usage can still be heard in a word that signifies the status of peasants under feudalism, ‘*Leibeigenschafft*’ (serfdom).

6. Towards the end of the 19th century several German comparative linguists and classical philologists who tried to reconstruct an Indo-European proto-language conjectured that originally the absolute use of \**es-* or of *εἶναι* was the *only* use. In his criticism of this developmental hypothesis Charles Kahn points out that already in Homer 80 % of the occurrences of this verb are copulative (Kahn 1973, 199 ff.).

7. The friend to whom I have dedicated this paper is sensitive to them.

*auch, wenn ich nicht mehr seyn werde, noch lieben und segnen. Unser Freund ist nicht mehr, er ist todt.*<sup>8</sup>

So Adelung, too, takes ‘*A ist*’ to be an ascription of actuality (primarily, of aliveness) to something. But his classification of this usage is the reverse of Bolzano’s: by his lights, the absolute use is one of many figurative (*figürlich*) uses of this verb most of which he regards as *elliptisch*.<sup>9</sup> The lexicographer may have been influenced here by a Wolffian philosopher. In a logic book that Bolzano often refers to Johann Peter Reusch maintained that sentences like ‘*Deus est*’ und ‘*ego sum*’ end with the copula: ‘here the predicate ‘*existens*’ has been dropped, for the sense of such sentences is: *Deus est existens, ego sum existens*’ (Reusch 383). Bolzano can hardly deny that ‘*A ist*’ is elliptical, for he maintains that sentences of this form have the same sense as ‘*A hat Daseyn*’ (*WL* II, 10, 64), and if that is correct, then ‘*A ist*’ must be elliptical for ‘*A ist daseyend*’, just as Reusch and Adelung claim.

Let us bracket the question which use of ‘is’, if any, is non-literal. The mere fact that the word ‘is’ is sometimes not used in the same way as in (S) is not a good reason for Bolzano’s preference. As we shall see shortly, the word ‘has’ is not only sometimes but very often used in other ways than in ( $\Sigma$ ).<sup>10</sup> All in all, the reasons Bolzano seems to have for preferring ( $\Sigma$ ) to (S) are not conclusive. But that should not prevent us from closely studying the logico-semantic features of the formulation he favours.<sup>11</sup>

If the word ‘has’ plays in a sentence the same role as in ( $\Sigma$ ), Bolzano calls it *Bindewort* or *Copula*. Here I want to reserve the title ‘copula’

---

8. Adelung 1793, art. ‘*Seyn*’, sub I, 2, no. (10).)

9. op. cit.

10. Elsewhere Bolzano contents himself with emphasizing the difference between the copulative and the absolute use of the verb: he points out that we are certainly not speaking of ‘actual’ objects when we say, ‘7 is prime’ or ‘What is impossible is not actual’. Cp. *WL* I, 191 sub 2, *WL* II, 11.

11. Of course, Bolzano did not overlook the obvious fact that very often even elementary predications contain neither ‘is’ nor ‘has’. He tries to break the resistance of sentences like ‘The sun shines’ in two steps. First he applies Aristotle’s *constructio periphrastica*: ‘Each inflected verb that is different from the word ‘is’ can be replaced, without any essential change of meaning, by the combination of ‘is’ with the (present tense) participle derived from that verb’ (1837, II, 10). According to Aristotle and Bolzano, ‘The sun shines’ is tantamount to (*durchaus gleichgeltend mit*) ‘The sun is shining’. (Try to forget about the English progressive aspect and to hear the second sentence as the English translation of ‘*Die Sonne ist scheinend*’.) In a second step Bolzano transposes this into his favourite key: ‘The sun has shiningness’. I think that even the first step in this procedure is dubious. For references to Aristotle and for a Fregean way of avoiding all these contortions see my (2006) 250ff.

for the word ‘is’ when used as in (S). So I need a different title for the word ‘has’ when used as in (Σ): henceforth I shall call it a connector.<sup>12</sup> Of course, ‘has’ is not in each of its modes of employment (as transitive verb<sup>13</sup>) a connector. Aristotle distinguished the following usages (among others): Socrates has a virtue ‘as a quality’, he has a snub-nose ‘as a part’, he has a house ‘as a possession’, and he has—a ‘very peculiar use of this word’, Aristotle remarks—a wife.<sup>14</sup> ‘Has’ functions as connector only if it is followed by a property designator (1837, I, 379f). Bolzano maintains that the sense of the connector is ‘wider than’ that of ‘has’ in any of its other modes of employment (loc. cit.). But in his technical acceptance of the term ‘wider than’, it is not defined for a concept like that expressed by the connector. A concept can only be wider than another concept if something falls under it (1837, I, 296ff, 543), but Bolzano explicitly says that nothing falls under the concept expressed by the connector (1837, I, 360, 558). That is a feature the connector shares with connectives like ‘and’ and ‘not’ (loc. cit.). Unlike the concepts *the Golem built by Rabbi Loew* or the concept *golden mountain*, the concepts expressed by the connector or by the connectives do not even have ‘the form of those that represent objects (*die Form derer, die sich auf Gegenstände beziehen*)’ (1837, III, 406),—only concepts expressible by singular or general terms, no matter whether they designate or apply to anything, have this form. So Bolzano’s ‘wider than’ talk concerning the connector concept must be non-technical in 1837, I, 379. The context makes it clear that he uses it to convey the following point: a snub-nose, a house and a wife are not *Beschaffenheiten*, but just as Socrates has the property of being courageous, so he has the properties of having a snub-nose, of having a house and of having a wife. The sense of the verb ‘to have’ in the specification of those properties is not the sense of the connector. In his ‘Größenlehre’ Bolzano puts the point as follows:

---

12. I have picked this term up from the admirable OUP translation of *WL* where it is used at two places (II, 9 and II, 505) to render Bolzano’s ‘*Bindewort*’. At the second place he calls ‘either—or—’ a *Bindewort*. As a matter of fact, in German grammar books this term is a standard title for connectives. By contrast, I use ‘connector’ exclusively for the word ‘has’ as used in (Σ) and all correct translations thereof.

13. To crown it all, we use this word also as an auxiliary verb to build the perfect tense of verbs (including that of the transitive verb ‘to have’).

14. Quotations from Aristotle’s dictionary article on ‘ἔχειν’ in *Cat.* 15, 15b17ff.; cp. also *Met.* Δ 23.

## TWO OPTIONS

*Alles, wovon [sc. zu Recht] gesagt werden kann, daß ein gewisser Gegenstand es habe, das oder dessen Haben ist eine Beschaffenheit dieses Gegenstandes.*

Everything of which one can [correctly] say that a certain object has it is such that either it is itself a property or having it is a property. (1833-, 98f.).

Since the word ‘has’ in the ‘that’-clause is ambiguous, I prefer a metalinguistic characterization of the two options:

If a sentence that expresses a truth about an object A consists of a term that refers to A followed by ‘has’ followed by a singular term *t* then either *t* itself or ‘the property of having’ followed by *t* refers to a property of A.

In Part 4 of this paper I shall argue that Bolzano should have opted for the second alternative more often than he did.

A major difference between the connector of Bolzano’s logic and the English word ‘has’ in ( $\Sigma$ ) will not matter in what follows, but it should not be swept under the carpet: the English word is tensed, and it can be pluralized, whereas Bolzano’s connector can no more be conjugated than the ‘ $\in$ ’ of set theory (1837, I, 202; II, 14f., 239).

Unlike the copula Bolzano’s connector is a dyadic predicator, and consequently what *he* calls the predicate term in ( $\Sigma$ ) is a *singular* term, whereas the predicate terms of traditional syllogistics are always general terms. To forestall confusion, I call expressions that can be obtained from a sentence by deleting *n* singular terms *n*-adic *predicators*,—they are what *Frege* calls predicates.<sup>15</sup> I classify as general terms expressions that can be transformed into a predicator by prefixing the copula, so, for example, ‘courageous’ and ‘(a) philosopher’ are general terms.<sup>16</sup>

Under a generous reading of ‘nominalization’ not only the abstract noun ‘courage’ in ( $\Sigma$ ) but also the quasi-description in

---

15. Apart from the dyadic predicator ‘*x* has *y*’ ( $\Sigma$ ) contains two monadic predicators, ‘*x* has courage’ and ‘Socrates has *y*’.

16. In an extended sense a copula is a linguistic element that transforms a general term or the infinitive of a verb into a predicator. Thus understood, not only the second word in (S) is a copula, but also the verb ending ‘-s’ in ‘The Earth moves’, and there is a copula of the second kind also in ( $\Sigma$ ). Cp. my (2006) 249–53. In ‘Socrates is Plato’s teacher’ the ‘is’ is followed by a singular term, but here this word is not a copula but a dyadic predicator: it can be replaced *salvo sensu* by the phrase ‘is identical with’ that contains a copula (cp. Frege 1892, 194). Bolzano discusses identity statements in 1837, II, 60.

( $\Sigma^*$ ) Socrates has the property of being courageous

is a nominalization of the adjective ‘courageous’. (I call it a quasi-description because it isn’t a definite description in the Russellian sense: if it were then the term ‘*a* property of being courageous’ should make sense, which it doesn’t.) Semantically, the quasi-description in ( $\Sigma^*$ ) shares with the abstract noun in ( $\Sigma$ ) the feature that you cannot understand it without *eo ipso* knowing what it designates. This does not hold of the definite description that follows the connector in

(D) Socrates has the virtue that was the topic of his debate with General Laches.

As a brief look at Plato’s dialogue shows, this description also designates courage.

Sometimes what follows the connector in the vernacular isn’t a singular term at all. Consider

(G1) Socrates has a virtue.<sup>17</sup>

About propositions expressed by sentences like (G1) Bolzano *should* say that their predicate concept is not expressed by what follows the connector. After all, this fragment of (G1) is a compressed quantifier phrase: (G1) comes to the same thing as ‘There is at least one virtue that (Socrates has)’. If we understand this along the lines of Bolzano’s account of ‘there is’ sentences (1837, II, 52–54), (G1) says about a concept that it is not empty. In Bolzano’s language,

(G1\*) The concept *a virtue of Socrates* has objectuality (*Gegenständlichkeit*).

Hence the predicate concept of the proposition expressed by (G1) is the concept *objectuality*.

A moment ago I declared that this is what Bolzano *should* say. Unfortunately, he is sometimes rather careless when it comes to characterizing

---

17. For reasons that I shall spell out in Part 4 of this paper, I believe that the ‘has’ in the structurally similar example ‘A has *a* sensation’ that Bolzano gives in 1837, II, 12/13 is not the connector.

what stands on the right-hand side of the connector. In 1837, II, 40 he says about the proposition expressed by

(G2) Caius has all virtues of a good father

that its predicate concept (*Aussagetheil*) is the idea of a totality (*die Vorstellung einer Allheit*). If (if!) this is correct, then in using (G2) we ascribe to Caius a property that is a totality of certain virtues. (One of them is thrift, as Bolzano claims on the same page. He seems to be thinking of his own Dad, Bernard Pompeius.) Suppose being  $B_1$ , being  $B_2$  ... and being  $B_n$  are all the virtues of a good father. Then the proposition expressed by (G2) is true if, and only if,

(H) Caius has the property of being  $B_1, B_2$  ... and  $B_n$ .

In (H) the connector is followed by a *singular* term that designates a 'conjunctive' property. But surely sentence (G2) does not have the same sense as (H). First of all, (H) does not even entail (G2): in order to obtain a logically equivalent formulation we would have to add: 'and no further property is a virtue of a good father'. Furthermore, the proposition expressed by (G2) does not contain the concept of any particular virtue, say that of thriftiness, whereas many such concepts are part of the proposition expressed by (H). However, (G2) can be paraphrased as follows:

(G2\*) Every *virtue of a good father* has the property of being had by Caius.

Bolzano should concede that the subject concept of the proposition expressed by (G2) is by no means identical with that of the proposition that Caius has the property of being thrifty: its subject concept is expressed by the italicized general term in (G2\*). What is represented by the predicate concept of this proposition is not a totality of virtues but a property that all virtues of a good father have. When Bolzano contends that the predicate concept of the proposition expressed by (G2) is the idea of a totality he is misled by the surface structure of the sentence. Many chapters of his *Philosophical Grammar*, the *Tractatus exponibilium* in 1837, II, 211–245, show that he tried very hard not to fall victim to this danger.

In spite of Bolzano's occasional remarks to the contrary, I claim that in his regimented version of the vernacular the position of 'b' in sentences of the form 'A has b' is reserved for *singular* terms.

2.

Some truth candidates, e.g. ( $\Sigma$ ), are about actual objects (*wirkliche Gegenstände*). An object is actual in Bolzano's (as well as in Frege's) acceptance of this term if, and only if, it is *wirksam*, that is, iff it is an element of the causal order.<sup>18</sup> Numbers and propositions are not actual: they are causally inert entities, unsurpassable do-nothings. Like every object they have properties, of course: some numbers are even, some propositions are true. Now what does Bolzano understand by '*Beschaffenheit eines wirklichen Gegenstandes* (property of an actual object)? Unfortunately, his usage in *WL* is unstable in a philosophically dangerous way. Sometimes he means by this phrase what he meant by it in his 1827 book 'Athanasia', that is, sometimes he takes properties of actual objects to be *adherences*.<sup>19</sup> In order to grasp this concept we must turn to the most important distinction in his ontology of actual objects that is to be found right at the beginning of his *Méditations Leibniziennes*:

**DIVISION**

B[1] Every object that is actual ... belongs to one of the following two kinds: either it exists in something else as one of its properties (*an etwas Anderem, als Beschaffenheit desselben*) or it is not a mere property belonging to something else but exists by itself (*für sich*), as one usually says ...

B[2] Philosophers usually call actual objects of the former kind, using a Latin word, adherences (*Adhärenzen*) and those of the latter kind substances. (1838, 21)

Entities of the former kind are, for example, the colour of a blossom and the sharpness of a blade. Claimants to the latter title are, for example, living beings, artefacts and rocks. As a teenager Bolzano read Baumgarten's 'Metaphysics', and there he must have come across a passage that largely anticipates what he says in B[1] and B[2]:

A[1] A thing can either not be actual (*würklich*) unless it is actual as a determination (*Bestimmung*) of another thing, or it can be actual although it is not actual as a determination of another thing.

A[2] The former is an accident (*Accidenz*) (something *cuius esse est inesse* [sth whose being is being in sth else]), and its actuality ... is being in virtue of something else (*inhaerentia*). The latter is a substance, something that exists

18. Cp. 1837 I, 362, 366, III, 16; 1838, 85; 1839, 2; Frege 1918, 76—and the Devil's question in Thomas Mann, *Dr. Faustus*, Kap. XXV: 'Ist wirklich nicht, was wirkt?'

19. Cp., e.g., 1837 I, 557; III, 10.

by itself, and its actuality is being-by-itself (*Vorsichbestehen, subsistentia*). (Alexander Baumgarten 1766, § 127.<sup>20</sup>)

The terminological divergence is as remarkable as the conceptual similarity. Baumgarten is obviously not one of those philosophers (alluded to in B[2]) who call actual objects that are not substances *adherences*.<sup>21</sup> Of course, Bolzano's out-of-the-way term rhymes with 'inherence', a term of the Aristotelico-scholastic tradition that was still used by Baumgarten and by Kant. But 'inherence' is a *singulare-tantum*: in A[2] it denotes an ontological status and Kant seems to use it to refer to a relation.<sup>22</sup> Bolzano changes the prepositional prefix, presumably because he likes the metaphor of adhering, or sticking, to something,<sup>23</sup> and because in German it sounds far more natural to say of the colour of a red blossom that it is something '*an der Blüte*' than to say that it is '*in der Blüte*'. He says only of mental adherences, e.g. acts of judging, ideas and sensations, that they are 'in' a subject (e.g. 1837, II, 67–69). More importantly, Bolzano employs 'adherence' as a *general* term that applies to the entities in the domain of the relation *x adheres to y*.<sup>24</sup> For centuries these entities were called *accidentia*,

20. A[1] *Ein Ding kan entweder nicht anders wirklich seyn, als wenn es in andern wirklich ist, das ist, als wenn es als eine Bestimmung eines andern Dinges wirklich ist; oder es kan auch wirklich seyn, wenn es gleich nicht als eine Bestimmung eines andern Dinges wirklich ist.* A[2] *Das erste ist ein Accidenz (... cuius esse est inesse), und seine Wirklichkeit, in soferne es ein Accidenz ist, ist ein Bestehen durch etwas anderes (inhaerentia). Das letzte ist eine Substanz, ein vor sich bestehendes Ding, und das Vorsichbestehen (subsistentia) ist die Wirklichkeit einer Substanz, insoferne sie eine Substanz ist.* (Baumgarten 1766, § 127.) The Latin phrase in A[2] is of Scholastic origin: around 1250, Boetius of Dacia (Bo of Denmark) used it in his commentary on Aristotle's 'Topics': '*Natura accidentis est, quod subiecto inhaeret* (The nature of an accident consists in inhering in a subject) ... *Accidentis esse est inesse.*' (Quoted after Ruedi Imbach, *Quodlibeta, Ausgewählte Artikel*, Fribourg 1996, 268.)

21. I did not find the term in any of the Leibniz-Wolffian and Kantian books on metaphysics that Bolzano frequently referred to. But two fairly remote authors he read use '*ein* [!] *Adhärenz*' or '*ein Adhärens*' as antonym of '*eine Substanz*': Joh. Heinrich Gottlieb Heusinger, *Ueber das idealistisch-atheistische System des Herrn Professor Fichte in Jena*, Dresden u. Gotha 1799, at p. 78, and Christian Friedrich Sintenis, *Pisteuon. Oder Ueber das Daysein Gottes*, Leipzig 1800, <sup>3</sup>1809 at p. 61. The second book stood on Bolzano's bookshelves, and he mentions it in 2005, 279.

22. The pair '*Inhärenz und Subsistenz*' appears in Kant's Table of Categories (1787, B 106) as first category of relation. But cp. B 229f.

23. Lucretius uses '*adhaerere*' in *De rerum natura* VI, 914, when he says about iron filings that they *stick to* a magnet. Bolzano uses '*haften an*' metaphorically, e.g., in 1838, 23–27, 284.

24. If German-speaking philosophers were to follow the very reasonable injunctions of the *Duden*, the 'official' dictionary of the German language, then what Bolzano calls *eine Adhärenz* would be called *ein Adhärens*,—'*veraltet für: Anhaftendes* (archaic for: sth attached to sth)', and the proper plural form would be '*Adhärenzien*'. The *Duden* knows the word '*Adhärenz*' only as *singulare-tantum*, '*veraltet für: Anhänglichkeit* (attachedness)'.

as they are still called in A[2] and in Kant's first *Kritik*.<sup>25</sup> Bolzano is aware of this usage, for when he employs the unusual term for the first time in his diaries of 1807/08, he writes: '*Adhärenz oder Akzidenz*' (2009, 128), and when he scrutinizes Kant's Table of Categories in his *WL* he replaces Kant's '*accidens*' without blinking an eye by the term he himself prefers (1837, I, 560, 562, 563; IV, 346). Why does he prefer 'adherence'? I think he wants to block the idea of contingency that tends to be triggered by the word 'accident'. After all, Aristotle pointed out that the word '*συμβεβηκός*' (that Boëthius rendered by '*accidens*') in one of its usages means as much as 'something that belongs to an object but not necessarily'.<sup>26</sup> Now this is not true of all adherences: according to Bolzano, God's kindness is an adherence that necessarily belongs to Him, while the kindness of Martin de Tours who used his sword to cut his cloak in two, and gave half of it to a scantily clad beggar, did not necessarily belong to that man (cp. 1837, I, 522; 1838, 283).

If an object has a certain adherence necessarily, then it has it always, but the inversion does not hold: 'For example, illness might well be a permanent property of a certain man'—Bolzano knows only too well what he is talking about—'without being a necessary one' (1837, I, 522). Some adherences belong only temporarily to their bearers.<sup>27</sup> A blossom that was first bright red and that is now dark red has changed. Changes are elements of the causal order: John's rude remark might cause Anne's blushing, and Anne's blushing might make John ashamed of himself. Substances are not changes,—they undergo changes. So the division of all elements of the causal order into substances and adherences would not be exhaustive if we were not allowed to classify changes as adherences. In his Philosophical Diary of 1807/08 Bolzano wrote: 'Substances can change ... These changes are accidents' (2009, 128). In *WL* he emphasizes that he takes the word *Beschaffenheit* 'in that wide sense in which ... every fleeting change of an

---

25. Husserl called them *individuelle Momente*. Nowadays they suffer from being called 'tropes' by Anglophone ontologists who seem not to care that for a couple of centuries this title was reserved for figures of speech. As was to be expected, German-speaking philosophers followed suit, so nowadays there are lots of publications on *Tropen* that are neither about figures of speech nor about tropical countries.

26. Aristotle, *Met.* Δ 30, 1025a14f.; cp. *Top.* A 5, 102b6.

27. According to Porphyry's very influential introduction to Aristotle's logical writings, it is true of every accident that it resides only for a while in its subject (*ὑποκείμενον*) and that its disappearance does not endanger the existence of the subject. Bolzano quotes the pertinent passage in Greek in 1837, I, 554.

object is at least a temporary property of it'.<sup>28</sup> Now I dare say that there neither is nor ever was such a wide sense of '*Beschaffenheit*'. Changes, no matter whether fleeting or taking ages, are events or processes, things that happen, take place, come to pass or go on, but it simply makes no sense to ask when a certain property happens etc. Bolzano is a philosopher prone to criticize other philosophers for deviation from well-established usage, so I take it that he would be ready to accept this kind of criticism. Instead of overstressing 'property', I prefer to distinguish two kinds of adherences: static ones and processual ones. Both the girl's blushing and the red of her lips are adherences, the former is a processual adherence (alias change), the latter is a static adherence. Static adherences may be ephemeral, relatively permanent or eternal, but they are not things that take place or happen. As Bolzano sometimes puts it, they are something that 'resides in certain things' rather than something that 'happens to something'.<sup>29</sup>

In 'Athanasia' Bolzano occasionally speaks of '*Beschaffenheiten oder Veränderungen* (properties or changes)', and he says of both that they 'stick to (*haften an*)' their bearers (1838, 24, 33f., 299f.). So here he does not misclassify changes as a special kind of properties, and he takes both to be adherences. In that disjunctive formulation '*Beschaffenheit*' apparently has a narrower meaning than in (DIVISION), and the disjunction seems to come to the same thing as 'static or processual adherences' in my terminology. Of course, the sense of the generic term 'adherence' still awaits elucidation.

### 3.

Bolzano explains the concept of a substance as follows:<sup>30</sup>

---

28. 1837, I, 379; vgl. 1837, III, 10. Bolzano's use of 'change' and 'state' is very confusing. He maintains that an ephemeral state of an object *is* a change (1837, III, 47). That category mistake certainly does not follow from the truth that an object must change if one of its states is ephemeral. He makes the very odd claim that every change only goes on for a short time (1837, III, 89, 500), although the melting of the pole caps is a change that takes a fairly long time. Even his assumption that every change has a temporal extension (1837, III, 96, 483) is not beyond doubt.

29. Cp. 1838, 26: *an etwas vorhanden sein* vs. *mit etwas vorgehen*.

30. 1838, 293 (cp. 445); 1837, IV, 185 (with reference to 1837, II, 65); 1851, 22. Bolzano's definition(s) of substance is (are) not my topic in this paper. A very thorough examination of this topic is to be found in Schnieder 2002, 171–243.

(Df. **Sub**)  $x$  is a substance  $:\leftrightarrow x$  is actual &  $\neg x$  is an adherence.

Tellingly, the hedged ‘affirmative’ clause of Bolzano’s introductory characterization of substances in part B[1] of (DIVISION), sc. ‘it exists by itself, as one usually says’, does not recur in his definition. That characterization was indeed ‘usually’ given by philosophers (recall Baumgarten’s ‘*Vorsich-bestehen*’ in A[2]), but, as Bolzano had already told his students in Prague, one ‘gives an incorrect explanation of the concept of a substance if one says that a substance is something that exists by itself’ (1834, III, 256). One good reason for rejecting such an explanation is this: it is not only the case that just as there cannot be adherences without substances there cannot be substances without adherences,<sup>31</sup>—the existence of a particular substance  $s$  can also be dependent on the existence of a particular adherence  $a$ , for some adherences belong necessarily to their bearers.<sup>32</sup> If such a situation obtains then neither  $a$  nor  $s$  can exist without the other,—neither of them can ‘exist by itself’. Hence not even the substance that Bolzano calls God is something that exists by itself, for He has each of his adherences necessarily. (So Descartes was wrong when he said of the definiens of his first definition of ‘substance’, i.e. ‘ $x$  does not stand in need of any other thing in order to exist (*nulla alia re indigeat ad existendum*)’, that it holds at least of God.<sup>33</sup>)

Bolzano’s purely negative explanation assigns conceptual priority to the concept *adherence*. Obviously, we would run into a circle, he notes, if we now defined an adherence as an actual object that is not a substance (1837, IV, 553). But he takes this concept to be definable (1837, IV, 346): adherences, he says, are ‘actual objects that occur only as properties in some other actual thing (*Wirklichkeiten, die sich bloß als Beschaffenheiten an einem andern Wirklichen befinden*)’ (1837, III, 10, cp. 1838, 22). In a

31. Some ‘modern philosophers’, Bolzano complains, make the ‘coarse’ assumption that the notion of a substance is the notion of something that remains when one has taken all its properties away; of course, they are right in maintaining: ‘*Ein solches Etwas gibt es u. kann es ... nicht geben*’ (1840–41, 121). As Ettore Casari once put it, tongue in cheek, there are no ‘Musilian substances’. Leibniz 1691, 364 turned this point against the 2<sup>nd</sup> definition of ‘substance’ in Descartes 1644, 25. Descartes thought that this term cannot be applied univocally to God and to other things, and according to his 2<sup>nd</sup> df. something is a substance if, and only if, it requires only God’s assistance (*solo Dei concursu*).

32. Cp. Simons 1987, ch. 8, on the difference between generic and individual dependence.

33. Descartes (loc. cit.). Leibniz continues (loc. cit.) that often (*saepe*) there is no particular accident on which the existence of a particular substance depends. So he reckons with the possibility that *sometimes* such a dependence does obtain, and he could have turned this point against Descartes’ 1<sup>st</sup> definition.

*different* object: no adherence is an adherence of itself. In an *actual* object: the evenness of the successor of three is a property of this number, but not an adherence, and the falsity of what the witness said is a property of a proposition but also not an adherence. *Faitnéants* like numbers and propositions have properties, but nothing adheres to them. We can codify Bolzano's definition of 'adherence' as follows:

(Df. Adh)<sub>B</sub>  $x$  is an adherence  $:\leftrightarrow$   
 $x$  is actual &  $\exists y (x \neq y \ \& \ y$  is actual &  $x$  occurs as a property  
in  $y$ ).

I prefer a more laconic definition that studiously avoids the word 'property'. I already gave you *one* of my reasons for this abstinence: changes are adherences, but unlike properties they happen or take place. (Other reasons for avoiding the use of 'property' in this context will follow.)

(Df. Adh)  $x$  is an adherence  $:\leftrightarrow \exists y (x$  adheres to  $y)$ .

This definition is just as little circular as

(Df. M)  $x$  is a murderer  $:\leftrightarrow \exists y (x$  murdered  $y)$ .

Soon I shall try to capture the conditions mentioned in Bolzano's verbose definiens by adding (under the name 'adherence principles') meaning postulates for the two-place predicator ' $x$  adheres to  $y$ '. This is a predicate that is indefinable also by Bolzano's lights. Already in 1812 he wrote in his Philosophical Diary: 'Currently it seems to me that the concept »occurring in something« is simple (*Es kommt mir jetzo vor: daß der Begriff des Seyns an etwas ... ein einfacher sey*)' (1981, 31). Five years later the diary writer repeats this conjecture with greater confidence: 'The concept »occurring in some other object« is a simple concept (*der Begriff des 'An einem Andern seyn' ist ein einfacher*)' (1981, 165). (In both passages, he identifies this concept with the concept expressed by the connector.<sup>34</sup> That move is highly problematic, as we shall see shortly.)

Let me now try to throw some light on the unanalysable sense of the predicator ' $x$  adheres to  $y$ '. Bolzano's definition of 'adherence' immedi-

---

34. In the groping reflections on '*Adhären*' and '*Haben*' that he wrote down in 1818/19 he also regards both verbs as stylistic variants (1980, 44f.).

ately implies (1) that the relation signified by ‘*x* adheres to *y*’ obtains only between *actual* objects and (2) that it is *irreflexive*. We can capture this by two meaning postulates:

(APr-1)  $\forall x\forall y (x \text{ adheres to } y \rightarrow (x \text{ is actual} \ \& \ y \text{ is actual}))$

(APr-2)  $\forall x\forall y (x \text{ adheres to } y \rightarrow x \neq y).$

(Each of the (APr-) conditionals begins with an operator that I suppressed: ‘It is a conceptual truth that’.) The relation of adhering is also *asymmetrical* and *intransitive*.<sup>35</sup>

(APr-3)  $\forall x\forall y (x \text{ adheres to } y \rightarrow \neg (y \text{ adheres to } x))$

(APr-4)  $\forall x\forall y\forall z ((x \text{ adheres to } y \ \& \ y \text{ adheres to } z) \rightarrow \neg (x \text{ adheres to } z)).$

Bolzano takes the satisfiability of the antecedent of (APr-4) for granted when he says that not only substances but also adherences are bearers of adherences (1838, 22f., 284). (If this is to hold of *every* adherence then every adherence of a substance is the first in an infinite series of higher-order adherences. As far as I know, Bolzano only gives examples of first- and second-order adherences.) One might get the impression that he declares adhering to be transitive when he writes that ‘from a certain point of view (*unter einem gewissen Gesichtspunct*)’ every adherence of an adherence can be conceived of as an adherence of the substance in which the latter occurs (1838, 22f.). But this impression of transitivity is deceptive. Suppose the blossom of the rose I am pointing at is bright red. Then the *brightness* of the red of this blossom is an adherence of that colour and not (as it would have to be if transitivity were to reign) of the blossom. When we take the ‘point of view’ Bolzano is talking about, what happens linguistically is this: the adjective ‘bright’ that applies to a colour is turned into what might be called the ad-adjective ‘bright-’ that combines with the adjective ‘red’ to form a term that applies to a coloured object. Or consider a processual adherence. If the ballerina is moving gracefully then the grace(fulness) of her movement is not one of *her* adherences. When we take the ‘point of view’ Bolzano is talking about, what happens linguistically is this: the adjective ‘graceful’ that applies to a movement is turned into an

---

35. Each of these two properties implies irreflexivity. Nevertheless, I refrain from eliminating (APr-2) because it echoes Bolzano’s own definition.

adverb that combines with the verb ‘moves’ to form a term that applies to a dancer.

While the last three principles characterized adhering in terms of the logic of relations, the next three principles will employ the mereological concept ‘is a proper part’. Parts of material bodies are—like the material bodies they are parts of—among Bolzano’s examples for substances (1838, 27f., 446). By contrast, adherences are not parts of their bearers, no matter whether the bearer is a substance or an adherence. The *sharpness of a knife-blade* is not related to the blade as the point of the blade is related to it. The ballerina’s graceful movement has (successive) parts, but the *grace(fulness) of her movement* is none of its parts. In Peter Simons’ notation (‘ $\ll$ ’ for ‘is a proper part of’):

$$(APr-5) \quad \forall x \forall y (x \text{ adheres to } y \rightarrow \neg x \ll y).$$

Saying that an adherence ‘occurs in something *as a property*’<sup>36</sup> is Bolzano’s way of denying that it occurs in something *as a part*.

Surely we can reckon with Bolzano’s consent if we lay down that the relation between an adherence and its bearer is not a part-whole relation in the other direction either:

$$(APr-6) \quad \forall x \forall y (x \text{ adheres to } y \rightarrow \neg y \ll x).$$

The bearer of an adherence cannot be one of its parts,—a pale face is not a part of its paleness, nor is a graceful movement a part of its grace.

Two objects neither of which is part of the other may have a part in common: two string quartets may share the same viola player,—after all, viola players are comparatively rare. By contrast, adherences and their bearers never overlap mereologically:

$$(APr-7) \quad \forall x \forall y (x \text{ adheres to } y \rightarrow \neg \exists z (z \ll x \ \& \ z \ll y)).$$

The paleness of the right cheek of a man who grew pale with terror is a part of the paleness of his face, and his right cheek is a part of his face. But there is nothing that is a part of his face as well as of its paleness. The movement of a ballerina has ‘successive parts’, the ballerina has ‘co-existent parts’, as Bolzano would put it (39). But the ballerina and her movement

---

36. Cp. DIVISION B[1].

have no part in common.<sup>37</sup> We may assume that Bolzano would be prepared to endorse (APr-7), too.

The next two principles could be called *dependence* principles. The existence of an adherence stands and falls with that of its bearer:

$$(APr-8) \quad \forall x \forall y (x \text{ adheres to } y \rightarrow (\forall t)(x \text{ exists at } t \rightarrow y \text{ exists at } t)).$$

Once a Palladio villa is blown to smithereens its beauty has passed away, too. A ballerina's movements will not go on after her death, and the grace of the pirouette we are now seeing will not outlive the end of that pirouette. Bolzano makes this point explicitly with reference to adherences of substances when he says: 'each adherence presupposes the existence of a substance in which it occurs, as a condition of its own existence'.<sup>38</sup> And surely he would be prepared to add that the existence of a second-order adherence  $a_2$  presupposes the existence of the first-order adherence  $a_1$  to which it adheres—and that it ultimately presupposes the existence of the substance to which  $a_1$  adheres, even though, strictly speaking,  $a_2$  itself does not adhere to that (or any other) substance.

Adherences are identity-dependent on their bearers:

$$(APr-9) \quad \forall x \forall y \forall z ((x \text{ adheres to } y \ \& \ x \text{ adheres to } z) \rightarrow z = y).$$

The sharpness of this blade differs from the sharpness of that blade (even if both blades are equally sharp), and the grace of the movement the ballerina is making right now is not the same as the grace of any other movement (even if equally graceful). Bolzano gestured in this direction already in 1810 when he wrote in a manuscript on *Allgemeine Mathesis* that the virtue of person A cannot be identical with the virtue of another person B no matter how similar A's virtue may be to B's.<sup>39</sup> In a letter he wrote a quarter of a century later he says about the colour of a particular rose:

37. Two non-atomic objects may have no part in common and yet occupy the same spatial region for a while: the pirouette of a ballerina takes place exactly where she is located while pirouetting. Cp. Simons (1) 211.

38. [Jede Adhärenz setzt das Daseyn einer Substanz, an der sie sich befindet, als eine Bedingung zu ihrem eigenen Daseyn voraus. (1834, I, 183.) Cp. 1838, 321/322; 1851, 110; and for mental states and mental events: 1838, 26.

39. 1810b, 61 Anm. Cp. 1837, I, 217, III, 10ff., 109, 112 on the dependence of the identity of those adherences that are cognitive mental acts on the identity of the agents of those acts. (That is one of the many points on which Bolzano and Frege agree: cp. Frege 1918, 67f.)

This red (numerically the same) can occur in no second rose. The red that occurs in a second rose may be similar, very similar, but it cannot be identical with it, just because it isn't the same rose: for two roses two reds are required.<sup>40</sup>

Considered in the light of (APr-8) and (APr-9) Bolzano's contention that it is the *rose* that is the object the colour adheres to requires a charitable reading. 'A whole ... is a distinct object that is essentially different from its parts (*ein eigener, von seinen einzelnen Theilen wesentlich unterschiedener Gegenstand*) (1837, I, 381), hence no rose is identical with its blossom. So according to (APr-9), the red Bolzano referred to cannot adhere both to that flower and to its blossom. Furthermore, if the blossom is cut off from the rose, then the red that Bolzano referred to survives for a while the death of the flower. So according to (APr-8) it does not adhere to that flower. Both problems disappear if we say: strictly speaking, that red adheres to a part of the flower.<sup>41</sup> After all, it's the blossom of the rose that is red—and not its stalk, its thorns or its leaves. When Bolzano speaks of the red of that rose then this is with respect to the bearer of the adherence a *totum pro parte* (a figure of speech that we also employ when we criticize the American policy towards Iran although we have only the US policy in mind). Bolzano is very well aware of this: 'We take the liberty of assigning an [adherence] to a whole even though in actual fact it only resides in, or happens in, one or several of its parts, ... so this way of talking is figurative.'<sup>42</sup>

Bolzano's conception of adherences does not exclude that an adherence resides in a collection (*Inbegriff*) of substances.<sup>43</sup> By stating that 'some adherences can only be conceived of as occurring in several substances' (1838, 23; cp. 1851, 110) Bolzano seems to contradict something Leibniz had emphasized very strongly: 'It is impossible that one and the same individual accident occurs in two subjects.'<sup>44</sup> The disagreement

---

40. *Dieß Roth (numero idem) kann sich an keiner zweyten Rose finden. Das Roth, das sich an einer zweyten Rose findet, kann jenem ... gleich, sehr gleich kommen, aber dasselbe kann es nicht seyn, eben weil es nicht dieselbe Rose ist; zu zwey Rosen werden zwey Röhthen erfordert.* (1935, 32f.) This is not the first appearance of the redness of a particular rose in Bolzano's writings: '*Die Röhthe, welche gerade dieser Rose ... zukömmt, ... kann eben darum nie einer Andern ... zukommen... Diese Rose hat diese Röhthe.*' (1805–07 [?], 83).

41. Cp. 1837, I, 387f. The problem is carefully discussed in Schnieder 2004, 197–215.

42. *Wir erlauben uns[,] eine gewisse Beschaffenheit oder Veränderung oft dann schon dem Ganzen beizulegen, wenn sie im Grund doch nur in einem oder etlichen Theilen desselben vorhanden ist oder vor sich geht... [Diese] Art zu sprechen [ist] also uneigentlich.* (1838, 33f.)

43. On Bolzano's conception of *Inbegriffe* cp. Simons 1997.

44. Leibniz 1716, 400f., echoing Aquinas, c.1255, Sent. I, dist. 27, q. 1, a. 1, ad 2.

is only apparent, however, if Leibniz has a *distributive* reading of the predicator ‘ $x$  occurs in the two subjects  $y$  and  $z$ ’ in mind, that is, ‘ $x$  occurs in  $y$  &  $x$  occurs in  $z$ ’, while Bolzano gives it a *collective* reading: ‘ $x$  occurs in an object that is composed of  $y$  and  $z$ ’, in his own words:  $x$  is ‘an adherence that sticks to a collection of several substances ([*eine*] *an einem Inbegriff mehrer Substanzen haftende Adhärenz*)’ (1837, III, 497). This second reading is required by (APr-9). When Itzhak Perlman and Pinchas Zukerman play the ‘Duo for Violin and Viola’ K. 423, then this performance is a (processual) adherence of a pair of musicians. (This pair is as much an actual entity as the two musicians are.) To be sure, the adherence of this ensemble can only exist if two other (processual) adherences exist, namely Itzhak’s playing the violin part and Pinchas’s playing the viola part. But a performance of K. 423 is, as Bolzano says, an adherence that ‘can only be conceived of as occurring in’ a duo.<sup>45</sup> (Incidentally, *collections* of substances that must exist if such adherences are to exist have themselves to be counted as *substances* if the adherence/substance distinction is to provide an exhaustive division of the realm of all actual objects.<sup>46</sup>)

The meaning postulates for ‘ $x$  adheres to  $y$ ’ nicely dovetail with Aristotle’s conception of entities that are ‘in a subject (*ἐν ὑποκειμένῳ*)’: ‘By “in a subject” I mean what is in something, not as a part, and cannot exist separately from what it is in’ (*Cat.* 2, 1a24f.). Those principles blend in equally well with what Leibniz said about (individual) accidents: ‘Accidents cannot become detached and walk around outside of substances.’<sup>47</sup> An accident cannot be transmitted like a virus or transplanted like

---

45. Two caveats are worth adding. (1) Zukerman’s playing the viola part and a machine reproducing his earlier playing of the violin part can make up a kind of performance of K. 423, but even this hybrid performance is an adherence residing in an *Inbegriff* (that consists of a man and a machine). (2) If Perlman plays his part in a hotel room in London while Zukerman plays his in a hotel room in Rome, then it is not the case that at that time a performance of the Duo by P. and Z. takes place. So two processual *person*-adherences very much like those that jointly make up a performance of K. 423 might exist, although no processual *ensemble*-adherence that is such a performance exists. An adherence of the latter type only comes into existence a day later, say, when the two musicians play together in a concert hall in Paris.

46. For a discussion of additional problems that arise for the exhaustiveness claim in (DIVISION) and for an attempt to dissolve them see Schnieder 2002.

47. Leibniz 1714, §7. “Are you in pain, dear mother?” [asked Louisa]. “I think there is a pain somewhere in this room,” said Mrs. Gradgrind, “but I couldn’t positively say that I have got it.” After this strange speech, she lay silent for some time.’ (Charles Dickens, *Hard Times*, II/9.) One of the reasons why the utterance Mrs. Gradgrind makes on her deathbed is strange is that pains are adherences.

an organ: 'It is inconceivable that an accident passes from one subject to another.'<sup>48</sup>

The most famous Italian novel of the 19th century contains a satirical echo of the conception of adherences or *accidentia* I have tried to elucidate. In *I Promessi Sposi* (The Betrothed) Alessandro Manzoni tells of the pestilence that raged in Lombardia in 1630, and in the course of this story we learn that the wrong application of a metaphysical distinction can have lethal consequences:

On the very first rumour of pestilence, Don Ferrante was one of the most resolute, and ever afterwards one of the most persevering, in denying it, not indeed with loud clamours, like the people, but with arguments, of which, at least, no one could complain that they wanted concatenation. '*In rerum natura*,' he used to say, 'there are but two species of things, substances and accidents; and if I prove that the contagion cannot be either one or the other, I shall have proved that it does not exist - that it is a mere chimera.' [Don Ferrante first argues to his own satisfaction that the contagion cannot be any kind of substance. Then he turns to the alternative option:] 'It remains to be seen whether it can be an accident. Worse and worse. These gentlemen, the doctors, say that it is communicated from one body to another; for this is their strongest argument and the pretext for issuing so many useless orders. Now suppose it is an accident. Then it must be a transported accident (*un accidente trasportato*), two words quite at variance with each other; there being no plainer and more established fact in the whole of philosophy than this, that an accident cannot pass from one subject to another (*un accidente non può passar da un soggetto all'altro*) ...'

*His fetus*, that is to say, relying on these beautiful principles, Don Ferrante used no precautions against the pestilence. He became infected and lay down to die.<sup>49</sup>

It hardly needs saying that the premature death of this quixotic Aristotelian, far from demonstrating the untenability of the substance/adherence distinction, only shows that one does well to acknowledge the substantiality of the bacterium *Yersinia pestis*.

---

48. Leibniz 1704, 224. Kant (1786, A 131) repeats this claim: '*Accidentia non migrant e substantiis in substantias* (accidents do not migrate from one substance to another).'

49. Alessandro Manzoni, *I Promessi Sposi, Storia milanese del secolo XVII* (1827, rev. edn. 1840-42), ch. 37. Being a son of an Italian father, Bolzano knew this book, of course. In 1837 he wrote from his 'exile' in Southern Bohemia to a friend who was about to travel to Prague: 'I take the liberty to ask you to bring this parcel to my brother. It contains the famous novel by Manzoni: *I sposi promessi* [sic]. If you fancy I can also lend it to you.' (Bolzano 2006, 209.)

Let us return now to Bolzano's conception of predication and of truth. For the author of *Athanasia* (1827) adherences are a special kind of *Beschaffenheiten*, namely properties of actual objects. The author of *Wissenschaftslehre* (1837), too, understands by 'adherences' properties of actual objects.<sup>50</sup> But in his magnum opus Bolzano very often does not mean adherences when he is talking about properties of actual objects. That is glaringly obvious in his explanation of the concept of a 'common or shared property (*gemeinsame oder gemeinschaftliche Beschaffenheit*)'. As a paradigm case he presents the property of having hands as something shared by (unmaimed) human beings and apes (1837, I, 523; cp. III, 68). Since adherences are bearer-individuated according to (APr-9), a common property of many actual objects cannot be an adherence. It is what I got used to calling an *attribute* that is *instantiated* by many actual objects.<sup>51</sup> Attributes, unlike adherences, are not actual objects,—they are causally inert. Even an attribute that is a causal power does not itself cause anything: that is only done by objects that instantiate it or by events involving such objects. (The property of being magnetic does not attract any iron filings.) Attributes are *praedicabilia*—something that one can ascribe (correctly or incorrectly) to an (actual or non-actual) object. (Since this explication employs the pragmatic concept of a speaker's ascribing something to something, it is only provisional with respect to Bolzano's theory. Later on I shall replace it by one that is more Bolzanian, or objectivist, in spirit.) Being a satellite of the Earth is an attribute that can correctly be ascribed to more than one actual object, and being a prime number is an attribute that can be truly predicated of more than one non-actual object.

Being a *natural* satellite of the Earth and being an *even* prime number are also attributes although they cannot *correctly* be ascribed to more than one object. They are the kind of attributes that Bolzano has in mind when

---

50. Cp., e.g., 1837, I, 557 and III, 10. (The word 'adherence' occurs only seldom in the four volumes of *WL* and hardly at all in his later work.)

51. The terms '*Attribut*' and '*instantiieren*' (as elucidated above) are not part of Bolzano's philosophical vocabulary. Don't be misled by the fact that Rusnock & George decided to render '*Beschaffenheit*' by 'attribute' (rather than by 'property'). Understandably, they were unimpressed by my (prognostic) lamentation that this does not go well with my comments on Bolzano's views, but they will understand that I continue to deplore their choice. (In 1837, I, 378 ff. Bolzano distinguishes two kinds of *Beschaffenheiten*, namely, *Eigenschaften* and *Verhältnisse*. Rusnock & George render this by 'two kinds of attributes, namely properties and relations', whereas I prefer: 'two kinds of properties, namely qualities and relations'.)

he contrasts the concept of a common property with another notion: ‘If a certain property of an object  $\alpha$  does not belong to any other object ..., then I call it ... an absolutely ... peculiar or exclusive property of  $\alpha$  (*eine ihm schlechterdings ... eigenthümliche oder ausschließlich zukommende Beschaffenheit*)’ (1837, I, 523). According to Bolzano’s theism, God is an actual object, and omniscience (*Allwissenheit*) is an absolutely exclusive property of this object (524). Bolzano does not think that *every* property of God—even His less spectacular knowledge that two plus two equals four—has this feature. But if one takes the terms ‘property of an actual object’ and ‘adherence’ as co-extensive, then because of (APr-9) *every* property of an actual object is (so to speak) the private possession of its bearer. So in the present context Bolzano is thinking of omniscience as an attribute that, unlike many other attributes of God, is uniquely instantiated by Him.

The conception of properties of actual objects as attributes is presupposed in Bolzano’s explanation of various concepts. Thus he defines the concept of a *respect in which two objects differ* as the concept of a ‘property that belongs to one of them but not to the other’ (1837, I, 530; cp. III, 417). Now suppose the two objects are actual, and we conceive of their properties as adherences. Then because of (APr-9) *every* property that one of the two objects has is a property that the other lacks. But Bolzano emphasizes that it is ‘one of the most difficult questions’ whether two actual objects have to differ in at least one respect (1837, I, 532f.). These respects must be *attributes* if the affirmative answer is not to be an utterly trivial truth.

In the context of comparing his conception of truth with those of other philosophers Bolzano employs the notion of a property that is *shared* by several actual objects:

If two objects, Caius and Titus, have a common property, e.g., learnedness (erudition, *Gelehrtheit*), then we are not mistaken to subsume both of them under the same concept *a learned man* and thus to assert the propositions that Caius is learned and that Titus is learned... If however Caius and Titus have different properties, e.g., [one has] learnedness and [the other has] ignorance, and we ascribe (*beilegen*) ignorance to the learned man and learnedness to the ignorant one, then we are mistaken. (1837, I, 130f.)

Of course, Caius and Titus have different *adherences* even if they are both learned or both ignorant, for they are different persons.<sup>52</sup>

---

52. References to common properties of actual objects can be found in every volume of *WL*. One additional example may suffice. Judgements (acts of judging) are mental adherences,

Attributes are ascribables, *praedicabilia*, whereas adherences are not,—that is an Aristotelian insight. Socrates’s wisdom, conceived of as an adherence, can no more than good old Socrates himself ‘be predicated of a subject (καθ’ ὑποκειμένου λέγεσθαι)’.<sup>53</sup> Bolzano sometimes loses sight of this fact. In his ‘Athanasia’ (1838, 23f.) he repeatedly talks as if one could ascribe (*beilegen*) adherences. In *WL* we find a whole section about *Sätze, die eine psychische Erscheinung aussagen*, about ‘propositions that predicate a mental phenomenon’. This is my mercilessly literal rendering of the title of §143. In the OUP translation of *WL* the title sounds as harmless as can be: ‘Propositions *dealing with* psychological phenomena’. No doubt, some propositions do deal with mental phenomena, but what Bolzano maintains in this section is that in some propositions such phenomena

---

hence actual objects. That does not prevent Bolzano from writing a section with the title ‘Some properties common to all judgements (1837, III, 108ff.).—So I find Betti’s contention (2012, 175) that ‘there is no direct textual support for universalia [multiply instantiated attributes]’ in *WL* truly amazing. Not only are there dozens of passages in which Bolzano says of actual objects that they share properties,—there are even more passages in which he says this of non-actual objects. The property of not being actual, he asserts, is ‘the first quality that must be recognized to belong to all propositions’, and ‘a second property that all propositions have in common is ... that they are complex’ (1837, II, 4). ‘I call propositions similar if they have so many common properties that they are easily confused’ (1837, II, 92), etc. etc. On the other hand, there is not the slightest textual support for Betti’s claim that Bolzano regards all such statements as tantamount to assertions to the effect that two objects have ‘tropes’ that fall under the same concept. He never maintains that, say, the non-actuality of proposition P differs from the non-actuality of proposition Q because P and Q are different propositions. Betti tries to support her interpretation (op. cit. 172) by referring to Bolzano’s oddly phrased statement in 1837, I, 381 ‘that every distinct object also has its own properties (*daß jeder eigene* [?] *Gegenstand auch seine eigenen Beschaffenheiten habe*)’. But as the context shows, his point is not, as she claims, that all properties are ‘bearer-individuated’, but that any two distinct objects differ as to *some* of their properties. He concludes from that statement ‘that every whole has *certain* properties that do not also belong to its parts’ (loc. cit., my emphasis). Obviously this allows for the possibility that a given whole shares those properties with many objects that are not its parts. (The Royal Concertgebouw Orchestra can do what none of its members can do, namely perform Mahler’s Symphony of a Thousand, but it shares this property with several orchestras.) If Betti’s reading of the quoted statement were correct, Bolzano should have said that a whole has *no* property that also belongs to its parts.

53. *Cat.* 2, 1a20–b9. I follow the interpretation of this chapter that was accepted from Porphyry to Ackrill (even though it is no longer uncontroversial).—It is surprising that Aristotle seems to have lost sight of adherences in his formulaic characterization of substances in his *Metaphysics*: ‘A substance is a subject of which other things can be predicated though it cannot itself be predicated of a subject’ (*Met.* Δ 8, 1017b13–14, 23–24; Z 3, 1028b36–1029a2, 1029a8–9). This requirement is met not only by Socrates but also by his paleness, but this point plays no role whatsoever in Aristotle’s own reflections on the insufficiency of the *primum-subjectum* formula in *Met.* Z 3.

are what is predicated, and that is very doubtful. Bolzano classifies mental phenomena, e.g. sensations, as adherences. So what Bolzano in that section of *WL* refers to as *the sensation D*—let's take it to be the worst headache Anne ever suffered from—is allegedly predicated of Anne when we say:

( $\psi$ ) Anne has the sensation *D*.

Bolzano takes ( $\psi$ ) to be an instance of his schema 'A has b' in which a designator of an adherence occupies the 'b' position. By hyphenating ( $\psi$ ) in the manner of '*A—hat—die Empfindung D*' he takes care to make this point conspicuous. So according to §.143 of *WL* the 'has' in ( $\psi$ ) is the connector that also figures in 'Socrates has the property of being courageous'.

I think this is muddled. Bolzano should have chosen here the second of the TWO OPTIONS that he distinguishes in 'Größenlehre' (as quoted at the beginning of this paper): The property that is ascribed to Anne is not the worst headache of her life but rather the property of having (of undergoing) the worst headache of her life. Only the first occurrence of the verb 'to have' in 'Anne *has* the property of *having* the sensation *D*' is the connector,—just as only the first occurrence of that verb in 'Socrates *has* the property of *having* a snub-nose' is the connector. Anne's headache is just as little as Socrates's nose a *praedicabile*. What is predicated of Anne is not an adherence but a uniquely instantiated attribute.

Thus I regard the following sentence to be what treatises on figures of speech call a *zeugma*, a 'yoking together':

(Z) Anne has the property of being charming and the worst headache of her life.

In (Z) one verb is forced into two conflicting services, as is done with the verb 'to take' in Alexander Pope's description of Hampton Court as the place where Queen Anne does 'sometimes counsel take and sometimes tea'—and with the verb '*heißen*' in the German comedian's address to his audience, '*Ich heiße Heinz Erhardt und Sie willkommen*'. Sentence (Z) is a somewhat wooden example of the same genre.

There are several other occasions in *WL* where I think Bolzano would have done well to pick the second of the TWO OPTIONS. A harmless example is provided by his remark about the proposition that Caius has a brother who is not learned: 'the predicate concept of this proposition

contains the concept of negation (*in der Prädicatvorstellung dieses Satzes [an sich] ist der Begriff der Verneinung enthalten*)' (WL II, 270). This is true enough, but what *is* the predicate concept of this proposition? Bolzano does not pause to explain, so I suspect that he takes the answer to be obvious because of the surface structure of the sentence. But the expression following 'has', sc. 'a brother who is not learned' does not even purport to designate a property, hence it does not express a predicate concept. The property that Caius is said to *have* (in the connector sense of this word) is *the property of having a brother who is not learned*, and it is the italicized singular term that expresses the predicate concept of the proposition. My final example is more serious. Why does Bolzano take it to be a matter of course that the extension (*Umfang*) of a concept is a property of that concept (I 298, 543)? Presumably, because it is true to say that the concept *golden bracelet*, unlike the concept *golden mountain*, has an extension (in the Bolzanian sense of the word '*Umfang*').<sup>54</sup> But then, it is also true to say that Socrates has a snub-nose, which does not make that olfactory organ a quality of the man. So we need a better argument, and none is forthcoming. The (Bolzanian) extension of the concept *first human being who ever set his foot on the Moon* is Neil Armstrong, but the astronaut is certainly not a property of anything. What a non-empty concept *has* (in the connector sense of this word) is not its extension but the property of applying to such-and-such objects.

## 5.

Unsurprisingly, the ambiguity of Bolzano's talk of properties of actual objects goes together with an equally unrecognized ambiguity of his talk of having:<sup>55</sup>

$$\begin{aligned} (\text{Having}_{\mathcal{A}}) \quad & \forall x \forall y (x \text{ has}_{\mathcal{A}} y \leftrightarrow y \text{ adheres to } x) \\ (\text{Having}_{\mathcal{I}}) \quad & \forall x \forall y (x \text{ has}_{\mathcal{I}} y \leftrightarrow x \text{ instantiates } y). \end{aligned}$$

Having<sub>A</sub> is a relation whose domain comprises only actual objects and whose range comprises only adherences. As we saw in section 3, the first

54. What Bolzano understands by 'the extension of concept C' is what you specify when you give a true affirmative answer to the question, 'What falls under C?' (1837, I, 297ff.), so the answer 'Nothing.' is excluded.

55. Textor 1996, 65 and Morscher 2008, 136 f. gave voice to the same suspicion.

of these generalized biconditionals can be transformed into a definition of the concept *adherence*:

(Df. Adh)  $y$  is an adherence  $:\leftrightarrow \exists x (y \text{ adheres to } x)$ .

In *WL* Bolzano says that one of the ‘two intimately connected concepts’ of having and of a property is ‘absolutely simple’, and he conjectures (with little confidence) that it is the concept of having that is absolutely simple and contained ‘as a part’ in the concept of a property (1837, I, 380). As you can see in (Df. Adh), I take this conjecture to be (or rather: to become) entirely correct if one reads ‘having’ as ‘having<sub>A</sub>’ and if one replaces ‘property’ by ‘adherence’.

Having<sub>1</sub> alias instantiating is a relation whose domain comprises all objects and whose range comprises only attributes. Already with respect to the logical features a dyadic relation may possess, it differs sharply from adhering. While adhering is irreflexive (APr-2), instantiating is *partim-reflexive*. On the one hand, some attributes instantiate themselves, e.g. being self-identical, being an object and being an attribute,<sup>56</sup> so this relation is *not* irreflexive,—in symbols (‘I’ for ‘instantiates’):

(i)  $\neg \forall x \forall y (xIy \rightarrow x \neq y)$ .

On the other hand, Lessing’s Nathan instantiates wisdom without being identical with wisdom, hence this relation is not reflexive either:

(ii)  $\neg \forall x \forall y (xIy \rightarrow x = y)$ .

While adhering is asymmetrical (APr-3), instantiating is *partim-symmetrical*. On the one hand, objecthood instantiates self-identity, and vice versa. So this relation is *not* asymmetrical:

(iii)  $\neg \forall x \forall y (xIy \rightarrow \neg yIx)$ ,

---

56. The concepts *object* and *self-identical*, Bolzano maintains, are all-embracing (1837, I, 459). So the properties of being an object (self-identical), itself being an object (self-identical), are self-instantiating. ‘Every property’, he writes, ‘has ... the property of being a ... property’ (1837, II, 546). So the property of being a property, itself being a property, instantiates itself. (Russell’s paradox requires protection against contradiction-inducing cases but not a total ban on self-instantiation.)

On the other hand, Nathan instantiates wisdom but *not* vice versa, hence this relation is not symmetrical either:

$$(iv) \neg \forall x \forall y (xIy \rightarrow yIx).$$

Finally, while adhering is intransitive (APr-4), instantiating is *partim-transitive*. On the one hand, this relation is not intransitive:

$$(v) \neg \forall x \forall y \forall z ((xIy \ \& \ yIz) \rightarrow \neg xIz);$$

for wisdom instantiates self-identity (as everything does), self-identity instantiates objecthood, but wisdom, too, instantiates this attribute (as every object does). On the other hand, this relation is not transitive either:

$$(vi) \neg \forall x \forall y \forall z ((xIy \ \& \ yIz) \rightarrow xIz);$$

for Nathan instantiates wisdom, wisdom instantiates being a virtue, but the latter is an attribute that the wise Jew does *not* instantiate.

Bolzano comes close to maintaining the partim-transitivity of instantiation when he writes (in the course of his criticism of Kant's confused slogan '*Nota notae est etiam nota rei ipsius*')

The property of a mere property is not always also a property of the object to which the latter belongs (*die Beschaffenheit einer bloßen Beschaffenheit ist nicht immer eine Beschaffenheit auch des Gegenstandes, welchem die letztere zukommt*)... Thus every property of a thing has among other properties also the property of being a mere property, whereas the thing itself need not be a property, but may very well be a substance. (1837, II, 546, my emphasis.)

The phrase I underlined suggests that Bolzano concedes that *sometimes* the property of a property is a property of the object to which the latter belongs'. That is true if one regards the predicator 'y is a property of x' as equivalent with 'x instantiates y', and it is false if one understands that predicator in the sense of 'y adheres to x'.

If we transform the general statement (Having<sub>1</sub>) along the lines of (Df. Adh) into a definition of 'attribute', then we obtain an explanation that is Aristotelian in spirit:

$$(Df \text{ Attr})^{Arist.} \ x \text{ is an attribute} \ :\leftrightarrow \ \exists y (x \text{ instantiates } y).$$

I call this explanation Aristotelian in spirit because Aristotle maintained: ‘If everybody were healthy then health would exist but illness would not, and if everything were white there would be white(ness) but not black(ness)’.<sup>57</sup> If this definition is correct, then Bolzano’s conjecture that the concept *has* is contained ‘as a part’ in the concept *property* (1837, I, 380) is again confirmed, this time for the instantiation reading of ‘having’ and the attribute reading of ‘property’. Bolzano comes close to underwriting the ‘Aristotelian’ definition when he says: ‘as for the concept of a property one can maintain that it is, if not identical, at least equivalent to that of something that is had’ (1837, II, 14).<sup>58</sup> But as has been repeatedly pointed out, there are two passages in *WL* that are incompatible even with the weaker equivalence claim. According to these passages, there is such a thing as the property of being a golden mountain,—a property that no object has, and there is even such a thing as the property of being a round square,—a property that no object can have (1837, II, 333, 456).<sup>59</sup> (Obviously, such properties can only be attributes, since the concept of a bearerless adherence is inconsistent.) So the Bolzanian conception of attributes is ontologically very generous:<sup>60</sup> if the general term ‘B’ is meaningful (no matter whether it applies to any object at all), the quasi-description ‘the property of being B’ and the corresponding name ‘b’ designate an attribute.<sup>61</sup> (Before Russell, nobody suspected this generosity has to be restricted somehow.<sup>62</sup>)

57. Aristotle, *Cat.* 11, 14a7–10; cp. *An. Post.* II.7, 92b7ff.

58. [*Es läßt sich von dem Begriffe einer Beschaffenheit behaupten ... , daß er, wo nicht einerlei, doch gleichgeltend sey mit dem Begriffe eines Etwas, welches gehabt wird.* Casari played off this statement against the passages to be mentioned soon: Casari 2006, 192–197.

59. 1837, II, 333, 456. (The passage in 1837, I, 427 that Schnieder 2002, 81f. adds to those mentioned by Berg, Morscher and Textor is not, or not clearly, pertinent.)

60. Friends and foes of this conception get their say in Künne 2003, 53–56; 2006, 279ff. u. 2007, 348–52.

61. ‘Clearly, the lack of a property is itself a property,’ Bolzano maintains. (1837, II, 47). Suppose that Socrates is not pale at time *t*. Is his lack of paleness at *t* an adherence of Socrates? Then the universe of actual objects comprises at each moment as many particular absences of paleness as there are non-pale actual objects,—and so for every general term that at some time or other does not apply to a particular actual object. You need not share Quine’s predilection for desert landscapes in order to find such a conception of the universe of actual objects fairly unattractive. If one assumes that the properties designated by nominalizations of negative general terms are attributes, no such overpopulation of the universe of *actual* objects threatens. The world of non-actual objects is as hospitable to all those attributes as it is hospitable to all propositions in the series that begins thus: *p*, it is true that *p*, it is true that it is true that *p*, etc. *ad infinitum*.

62. Cp. Künne 2006, 279–283 (‘Limits of Abundance’).

If we take the generosity demanding passages more seriously than the statement that allows for (Df. Attr)<sup>Arist.</sup>, we need another explanation of the concept of an attribute. Or is this a wild goose chase? In his ‘Größenlehre’ Bolzano suspects that any attempt to give a definition (*Erklärung*) of this concept is doomed to failure and so he contents himself with offering the following elucidation (*Verständigung*):

- (Attr-of) Something is (what I call) a property of an object if, and only if, it can be predicated of that object in a true proposition the subject of which is that object.  
*[Ich nenne] alles, aber auch das allein Beschaffenheiten eines Gegenstandes ..., was sich von diesem Gegenstande in einem der Wahrheit gemäßen Satze, dessen Subject derselbe ist, aussagen läßt.*<sup>63</sup>

Actually, (Attr-of) could serve as a definition of the two-place predicator ‘*y* is an attribute of *x*’, but what we are after is a definition of ‘*x* is an attribute’. Perhaps Bolzano gave up too soon: by taking recourse to the concept of a *proposition* his elucidation paves the way for an explanation.<sup>64</sup> Bolzano himself emphasizes that the predicate concept in a proposition is always a *Beschaffenheitsvorstellung*. Nothing prevents him from endorsing the inversion of this statement as well. So let us first define the term ‘objective representation of an attribute’ (‘attribute concept’, for short) in terms of ‘proposition’ and ‘predicate concept’:

- (Df. AC) *y* is an attribute concept  $\Leftrightarrow$   
 $\exists z$  (*z* is a proposition & *y* is the predicate concept in *z*).

This preparatory move allows us to explain the concept of an attribute as follows:

63. 1833–, 98. Obviously, ‘subject’ means here: what the proposition is about (what falls under its subject concept, what its subject term stands for).

64. Textor may have been the first to recognize this: Textor 1996, 62, but the way he developed the basic idea was faulty: Schnieder 2002, 83ff. The definition proposed in Schnieder 2002, 96 contains a comment on its own vocabulary that specifies the intended reading of ambiguous words in the definiens by means of words that are admittedly themselves ambiguous though not in the same way. It would be nice if one could avoid such contortions. Textor 2003, 85f. tries to improve his first attempt at a definition with some borrowings from Strawson. It would be nice if one could remain closer home. The ‘propositionalist’ definition I am about to offer complies with these wishes. It is due to Morscher (2004, 381).

(Df. Attr)  $x$  is an attribute  $:\leftrightarrow \exists y$  ( $y$  is an attribute concept &  $x$  falls under  $y$ ).

The recurrence of the word ‘attribute’ on the right-hand side makes this look circular, but thanks to the ancillary definition of ‘attribute concept’ it isn’t, at least not for this reason.<sup>65</sup> It is worth noting that—unlike its ‘Aristotelian’ predecessor—(Df. Attr.) lends no support to Bolzano’s conjecture that the concept of having is contained ‘as a part’ in the concept of a property (1837, I, 380).

---

65. When I presented an excerpt of this paper in Prague in July 2014, one of my commentators raised four objections against this proposal. (Some of them can be found already in his Hamburg MA thesis, as I came to remember belatedly: Schnieder 2002, 86ff.) *Firstly*, Df. Attr is circular on Bolzano’s analysis of the notion of a predicate concept (1837, II, 9). *Reply*: Does Bolzano *analyse* the notion of a predicate concept when he says, ‘I allow myself to call that part of a proposition that purports to represent the property that the proposition ascribes to its object the *Prädicatorstellung* of the proposition?’ I think he only tries to elucidate that notion. (Here is a comparison. Bolzano accepts ‘A point is what limits a line’ [‘A moment is what limits a time-span’] as an elucidation of the concept *point* [*moment*] without regarding it as an analysis, for he takes the notion *line* [*time-span*] that is invoked in that elucidation to be analysable by recourse to the notion *point* [*moment*]. Cp. 1805–09, 99 f, 1810a, 55, , *WL I*, 369, BBGA 2A.11/1, 147.) In any case, the invocation of the notion of a property seems not to be required for an elucidation or analysis of the notion of a predicate-concept. Bolzano might as well have said: ‘the predicate-concept of a proposition P is that part of P that purports to represent what P ascribes to its object’, or better (since a proposition cannot engage in the activity of ascribing) ‘... what the object of P has according to P’. Of course, the phrase ‘what the object of a proposition has (instantiates)’ applies to properties, but it does not express the notion of a property, since uninstantiated properties are not what the object of any proposition has. *Secondly*, on Bolzano’s ‘liberal’ conception of a proposition (1837, II, 16f.), Df. Attr is extensionally incorrect. *Reply*: Bolzano was well advised never to rely on that conception, for it asks us to regard grammatically ill-formed strings like ‘Socrates has snow is white’ as expressions of a proposition. (Incidentally, the all too liberal conception is incompatible with the elucidation or analysis invoked in the first objection.) *Thirdly*, on the ‘strict’ conception (1837, I, 393) circularity threatens, for the notion of a proposition may involve the notion of an attribute. *Reply*: According to Bolzano, the notion of a proposition is simple, hence it contains no notion whatsoever. *Finally*, Df. Attr is not extensionally correct unless it is essential to every attribute-concept that only attributes fall under it, but some of the propositions expressible by (\*) ‘Socrates has whatever I am thinking of at this moment’ seem to show that this isn’t the case. *Reply*: The appearance is deceptive. What follows the ‘has’ in an utterance U of (\*) does not express an attribute-concept if in U it refers to, say, the most famous snub-nose in the history of philosophy: what Socrates has (instantiates) is not that olfactory organ but the property of having it (in one of the many non-connector senses of this verb).

## 6.

How are attributes individuated? Bolzano never explicitly confronted this question that Frege and Quine rightly took to be very important. How would he have answered it if he had considered it? There are (at least) four passages in *WL* that must be taken into account if one wants to make a well-educated guess.<sup>66</sup> ‘Clearly, the lack of a property is itself a property,’ Bolzano maintains, ‘and the lack of the lack makes the original property reappear’ (1837, II, 47). The second conjunct is a circuitous way of saying that the lack of the lack of property *b* is identical with *b*. Thus the property of not being not-round is identical with the property of being round. However, the concept *lack of lack of roundness* is not identical with the concept *roundness* (and the ‘concrete’ concept *not not-round* also differs from the concept *round*).<sup>67</sup> The next passage generalizes this remark:

If the two expressions

Whatever has *a*, has *b*

Whatever has *a*, has *c*

are really supposed to express two propositions, then the ideas *b* and *c* must be really different. It does not follow from this, however, that they must represent different properties. Rather, the above propositions would be not only different but also both true if *b* and *c* were only a pair of co-extensional ideas.<sup>68</sup>

So Bolzano quite clearly assumes that properties (attributes) are less finely individuated than concepts. I think we should applaud this stance, for a concept is a representational entity (*objective Vorstellung*), it is a potential content (*Stoff*, matter) of thinking, and different contents can represent one and the same object in different ways. By contrast, attributes are *repraesentanda*: they can only be intentional objects of cognitive acts and states, rather than their contents.

---

66. Schnieder 2002, 136 already drew attention to two of the four passages I shall analyse in what follows.

67. A concept of the form ‘not not-*B*’ is always different from the corresponding concept expressed by ‘*B*’, for the latter either does not contain the concept of negation at all or it contains it less often than the former (1837, I, 206, 447; II, 273).

68. „Wenn die zwei Ausdrücke ‘Was *a* hat, hat *b*’, ‘Was *a* hat, hat *c*’ wirklich zwei Sätze [an sich] vorstellen [ausdrücken] sollen: so müssen die Vorstellungen *b* und *c* wirklich verschieden seyn. Daraus folgt aber keineswegs, daß auch die Beschaffenheiten, welche sie vorstellen, verschieden seyn müssen. Vielmehr wären die obigen Sätze nicht nur verschieden voneinander, sondern auch beide wahr, wenn gleich die Vorstellungen *b* und *c* nur ein paar Wechselvorstellungen wären.“ (1837, II, 410f., Anführungszeichen eingefügt.)

And how *are* attributes individuated? The following proposal individuates them more finely than the extensions of the singular terms that can be used to ascribe them, and it is compatible with the passages I referred to:

(IAI) The attribute  $x$  is identical with the attribute  $y$  if, and only if, nothing can instantiate  $x$  without instantiating  $y$  and *vice versa*.

According to this principle of **I**ntensional **A**tttribute **I**ndividuation, being an equilateral triangle and being an equilateral triangle are one and the same attribute, whereas the concept *equilateral triangle* is not identical with the concept *equiangular triangle*, as Bolzano emphasizes (1837, I, 272, 274). According to (IAI), there is exactly one attribute that no object can instantiate, and exactly one that all objects must instantiate. Thus being a sword without a blade is identical with being a male vixen, and being an object is the same attribute as being not both coloured and colourless. Of course, the concepts *sword without a blade* and *male vixen* are different, and so are the concepts *object* and *not both coloured and colourless*,—after all, their makeup is as different as can be.<sup>69</sup>

Can Bolzano underwrite the intensional individuation of attributes? There is *prima facie* evidence to the contrary. In a passage I already had reason to appeal to he speaks of a ‘concept of a self-contradictory property, e.g., that of a round square (*Vorstellung einer sich widersprechenden Beschaffenheit, z. B. der eines runden Quadrats*)’ (1837, II, 333). By ‘property of a round square’ he does not mean a property that round squares have,<sup>70</sup> but the property of being a round square. The ‘*exempli gratia*’ in Bolzano’s formulation strongly suggests that he assumes that there is *more than one* ‘self-contradictory property’, and that assumption is incompatible with (IAI). But note that in the context from which I took this remark this assumption is pointless. What Bolzano needs is an example for a special kind of *concepts*. So he might as well have said: ‘the concept of a self-contradictory property, e.g., the concept of the property of being a round square’. There are plenty of such concepts, and the question whether there is more than one self-contradictory property is entirely irrelevant to the point at issue. So the formulation that speaks against (IAI) might be just a slip of the pen.

---

69. Pleas for (IAI) that are not concerned with questions of Bolzano exegesis can be found in Künne 2003, 26, 148f., 415f.; 2006, 263–266; 2007, 340–348 and in Schnieder 2004, 59–69.

70. The attribute of being round and the attribute of being square, according to Meinong(ians).

But there is another recalcitrant passage to be coped with. Starting from the observation that everything falls under the concept *object*, Bolzano wants to show that ‘there are several, indeed infinitely many (*mehre (ja selbst unendlich viele)*) concepts’ that are equally all-embracing (1837, I, 459). His first example is the concept *not nothing*<sup>71</sup> that contains the concept of negation twice. This concept is the first of an infinite series of co-extensional concepts ‘in which the concept of negation is repeated an even number of times (*die den Begriff der Verneinung in Wiederholungen nach einer geraden Zahl enthalten*)’. (A concept containing the concept of negation *n*-times is different from a concept that contains it less than *n*-times.) *Quod erat demonstrandum*, Bolzano could announce and move on to another topic. But instead he goes on to say: all-embracing concepts are ‘likewise all ideas of the form “something that has the property *b*” if *b* is a property that belongs to all objects without exception, such as the property of being self-identical, or the property of having only one of any two contradictory properties, etc.’<sup>72</sup> The phrase ‘such as ... or ...’ strongly suggests that he assumes there is *more than one* all-embracing property, and that assumption is incompatible with (IAI). But, once again, the context from which I took this remark renders the assumption pointless. He might as well have continued: All-embracing concepts are ‘likewise all concepts of the form “something that has the property *b*” if *b* is a property that belongs to all objects without exception, such as the concept something that has self-identity, or any concept of the same form as the concept *something that has the property of being either square or not square*’. There are many all-embracing concepts, and the question whether there is more than one all-embracing property is beside the point. Once again, the formulation that speaks against (IAI) might be just a slip of the pen.

So it seems to me that Bolzano has no good reason to *reject* the principle of Intensional Attribute Individuation.

---

71. I assume that by ‘*x* is not nothing’ Bolzano means: it is not the case that *x* is not an object.

72. ‘[I]ngleichen alle Vorstellungen von der Form: Etwas, das die Beschaffenheit *b* hat, wenn *b* eine Beschaffenheit ist, die einem jeden Gegenstand ohne Ausnahme zukommt, wie etwa die, mit sich selbst einerlei zu seyn, oder von je zwei widersprechenden Eigenschaften nur Eine zu haben u. s. w.’ (1837, I, 459).

Common properties—attributes that are multiply instantiated—are *universals* in Aristotle’s understanding: Universal (καθόλου) is something that is shared (κοινόν), i.e. something that belongs to (ὑπάρχειν) several objects.<sup>73</sup> Bolzano quotes this in *WL*, but instead of identifying universals with common properties he gathers that ‘the ancient logicians ... hardly meant by universals something different from what I call general concepts (*kaum etwas Anderes als was ich ... einen Gemeinbegriff nannte*)’, that is, concepts under which at least two objects fall (1837, I, 546). He is by no means certain of this: ‘I will not venture to decide’, he says, ‘whether universals were really supposed to be general concepts’ (1837, I, 224). He misinterprets the Aristotelian predicator ‘y belongs to x’ on p. 546 as ‘x falls under (the concept) y’, although on p. 123f. he had correctly understood it in the sense of ‘x has (the property) y’.<sup>74</sup>

For realists a universal is not a common concept, but they concede there is a close relationship between such concepts and universals: a universal b is a property such the concept of an object that instantiates b is a common concept. Some pages after the misinterpretation I have been complaining about Bolzano quotes a fairly long passage from the logic of the Thomist Antoine Goudin<sup>75</sup> without realizing that, in the first sentence, the Parisian theologian characterizes universals as common properties: ‘*Universale est unum aptum inesse multis*’ (Goudin 1671, 149). His concept of *inesse* is clearly not the concept of adhering to but that of being instantiated by. Now a few pages earlier in the same book, Goudin, whose ‘acuteness (*Scharfsinn*)’ Bolzano readily acknowledges, had described this kind of entity in more detail and distinguished it from another kind with which it can easily be confused (as Bolzano inadvertently shows): ‘An ontic universal is one that is capable of being in, or being predicated of, many objects, as, for example, human nature.’<sup>76</sup> Clearly, Goudin is here speaking of what

73. Cp. *Met. Z* 13: 1038b11–12.

74. The verb ‘κατηγορεῖσθαι’ (if understood in the sense of ‘to be *correctly* predicated of’) that is used by Aristotle in his almost homophonous characterization of the universal in *De Int.* 7, 17a39–40 is extensionally equivalent with ‘ὑπάρχειν’, for something is truly predicated of A iff it belongs to A. This something is not a concept but a property, as Bolzano had acknowledged in the earlier passage in *WL*.

75. 1837, I, 554f. Bolzano quotes Goudin (1671) also in 1837, I, 25 and 463, always misspelling his name as ‘Gaudin’.

76. [U]niversale in essendo est unum aptum inesse in multis, seu, unum aptum praedicari de multis, ut natura humana est in multis individuis, ac de illis dicitur (Goudin 136).

Bolzano calls a common or shared property. From a *universale in essendo* he carefully distinguishes a *universale in repraesentando*, i.e. ‘one thing that represents many (*unum multa repraesentans*),... such as an idea (*idea*) of a house in the mind of an architect that represents many houses’.<sup>77</sup> This idea is what Bolzano would call a *subjective Gemeinvorstellung*. The content of such an idea is what figured above as a common concept.

8.

At the very beginning of this paper I cut out the hard core of Bolzano’s conception of truth by a biconditional that applies only to a small set of truth candidates. I shall now turn to his decomposing explanation of the concept *true*.<sup>78</sup> Its best formulation is contained in a letter to Exner:

**(Df. T)** A proposition is true if, and only if, every object that falls under the subject concept of the proposition has a property that falls under its predicate concept.<sup>79</sup>

Here we have to keep in mind that, by Bolzano’s Aristotelian lights, propositions of the type ‘All S are P’ are false if no object falls under their subject concept. So the definiens of (Df. T) entails: There is at least one object that falls under the subject concept of the truth candidate. Sometimes exactly one object falls under the subject concept of a truth, sometimes several. So the definiens of (Df. T) is just as applicable to the proposition that all assassins of tyrants have courage, as it is applicable to the proposition that Socrates has courage. In the Philosophical Grammar that occupies large stretches of volume II of *WL*, Bolzano tries very hard to support his conjecture that each and every proposition can be

---

77. loc. cit.

78. The first paper exclusively devoted to this topic was Simons 1999, and as far as I know it is still the only paper of its kind. I have granted Bolzano’s conception of truth only a few brushstrokes on a large canvas: it appears in Künne 2003, ch. 3, 107–112 as one version of the ‘Classical Correspondence’ view of truth (as opposed to the ‘Cambridge Correspondence’ view). I still think this subsumption is both correct and important.

79. [*Ein Satz an sich ist genau dann wahr, wenn jeder Gegenstand, der der Subjectvorstellung des Satzes untersteht, eine Beschaffenheit [hat], die der Prädicatvorstellung desselben untersteht*] (Letter to Exner, 18/12/1834, in 1935, 90). I took the liberty of replacing Bolzano’s semantical metaphor ‘stand under (stehen unter)’ by the more familiar metaphor ‘fall under’ that has become the common coin ever since Frege used it.

expressed by a sentence instantiating the schema 'A has b' (if 'A' is taken to be a dummy both for singular terms and for phrases of the form 'every such-and-such').

If (Df. T) is correct then with respect to generality the predicate concept in a truth is in the same boat as the subject concept: in some truths exactly one property falls under the predicate concept, in some truths several properties do.<sup>80</sup> If the second state of affairs obtains, in Bolzano's words: 'if the predicate concept [of a proposition P] represents several properties (*sofern die Prädicativvorstellung der Beschaffenheiten mehre vorstellt*)' (1837, II, 26), then P is true, according to (Df. T), just in case every object that falls under the subject concept of P has at least one of those properties. Suppose a proposition is about exactly one actual object. Are the relevant properties adherences, or are they attributes? I shall call the first interpretive option the *Adherence Reading* and the second option the *Attribute Reading*.<sup>81</sup> Which of these two interpretations would, or should, find Bolzano's support?

For any truth of the form 'A has b' that is about one actual object there are as many adherences as there are actual objects that can correctly be said to be B. If the Adherence Reading is correct then the predicate concept of the truth expressed by ( $\Sigma$ ), 'Socrates has courage', is such that Socrates' courage, Jeanne d'Arc's courage, Rosa Luxemburg's courage and Sophie Scholl's courage are four of the many adherences that fall under it. According to the Adherence Reading, the proposition expressed by ( $\Sigma$ ) is true in the sense of (Df. T) if, and only if, at least one of these many adherences adheres to the object that falls under the subject concept of that proposition. Is that Bolzano's view?

In §.131 of *WL* Bolzano poses the question 'whether when we ascribe the property b to an object A we also ascribe to A all properties that fall

---

80. Some propositions are such that there is *no* property that falls under its predicate concept. Bolzano forgot to give us an example. Let me make good for the omission: (i) Suppose that on Good Friday 1830 at midnight Bolzano was sound asleep. Then the proposition expressed by the next sentence contains a predicate concept under which, as a matter of fact, no property falls: 'Anna Hoffmann has *the property that Bolzano ascribed to Anna on Good Friday 1830 at midnight*'. (ii) The proposition expressed by the next sentence contains a predicate concept under which no property *can* fall: 'Anna has *the property that Bolzano both had and lacked on Good Friday 1830 at midnight*'.

81. Schnieder 2002, 154f. pleaded for the Adherence Reading, and Arianna Betti 2012, sect. V, followed suit. Simons in his 1999, 16 (drawing attention to the Caius example in 1837, §.131) and in his 2006, 160 endorsed the Attribute Reading as a matter of course. I shall try to decide the issue in what follows.

under [the concept] *b'*, and he declares it obvious that the answer is negative. He employs an example to make this obvious to his readers,

(C) Caius has intelligence.

How does he proceed? If he were a proponent of the Adherence Reading, then surely he would have seized this optimal occasion for arguing: 'Look, it's perfectly clear that in making that statement about Caius we do not claim that he has *all* the adherences that fall under the concept of intelligence—including Aspasia's intelligence and Zeno's intelligence.' But how *does* Bolzano argue? He points out that in making that statement we surely do not commit ourselves to the claim that Caius has 'every kind of intelligence one can think of, for example, a well-developed, as well as a crude, a human, and an angelic intelligence (*englischer Verstand*), etc.' (1837, II 26).<sup>82</sup> No adherence is a *kind of* intelligence, but there is a kind of intelligence that is a common property of, say, all brilliant mathematicians. Kinds of intelligence are *attributes* that are multiply instantiable. If (Df. T) is correct then the proposition expressed by (C) is true if, and only if, Caius has (instantiates) at least one of the attributes that fall under the concept of intelligence. Bolzano might have made his point as well by using ( $\Sigma$ ) as his example. There are several kinds of courage, such as civil courage and military courage.<sup>83</sup> In stating that Socrates has courage we surely do not commit ourselves to the claim that Socrates has every kind of courage. Kinds of courage are multiply instantiable attributes. If (Df. T) is correct then the proposition that Socrates has courage is true iff Socrates has (instantiates) at least one of the attributes that fall under the concept of courage.

In his remarks on the 'concordance (*Einhelligkeit*)' of concepts Bolzano also presupposes that *attributes* of actual objects fall under the concept of a property. Two concepts are concordant just in case at least one object

---

82. Rusnock & George are aware of a historico-linguistic fact one has to explain to many German readers nowadays: '*englisch*' means here as much as '*den Engeln eigen*'. After all, *der englische Gruß* is not John Bull's 'How are you doing?' but the Angel's salute to Mary, 'Hail, thou that art highly favoured' (Luke 1: 28), and Goethe's Wilhelm Meister does not mean to say that Marianne looked like an English girl when he acknowledges with delight that 'in her white dressing gown' she really looked rather *englisch* ('*Wilhelm Meisters theatralische Sendung*', I/20). (As Bolzano's biographer *in spe* I report with regret that he would have stared with a frown at the latter use of the adjective.)

83. Cp. the sermon 'On courage' that Bolzano delivered as university chaplain on 25.02.1810, 264–270.

falls under them (1837, I, 440ff.). Thus the ‘concrete’ concepts *pious* and *learned* are concordant because there is somebody, e.g., St. Jerome, who is both pious and learned. But the corresponding ‘abstract’ concepts *piety* and *learnedness* are not concordant, for there is no property that is both a ‘kind of learnedness (*Art von Gelehrsamkeit*)’ and a ‘kind of piety (*Art von Frömmigkeit*)’. By contrast, he emphasizes the abstract concepts *prudence* and *wariness* are just as concordant as are the concrete concepts *prudent* and *wary*. Hence there should be a kind of prudence that is also a kind of wariness, and I think that thriftiness exactly fits that bill. (Bolzano does not give an example.) Now if Bolzano thought that it is adherences that fall under the concepts of human properties, he would have argued along the following lines: ‘There is no adherence that falls both under the concept *learnedness* and under the concept *piety*, whereas there is an adherence, e.g., my father’s thriftiness, that falls both under the concept *prudence* and under the concept *wariness*.’ But there is no trace of any such argument. Take any property *b* you like, no adherence is a *kind of b*. A special kind of prudence is a common property of many people.—Actually, I have never come across *any* passage in which Bolzano says of entities that fall under a property concept that they are adherences.

Let us see how Bolzano describes in §.131 the situation that arises ‘if the predicate concept of a proposition represents several properties’ (as he puts it at the beginning of that section):<sup>84</sup>

- (1) We must say ... that the proposition [that A has b] leaves entirely undecided which of the many properties that fall under its predicate concept (if there are more than one) belong to the one object, or to the several objects, that fall under the concept A.

---

84. Jan Berg has repeatedly maintained that the predicate concept of a true proposition is always a concept, ‘*dessen Extension genau eine Beschaffenheit umfaßt*’ (BBGA 2B.16/1, 164, note). But (Df. T) does not speak of ‘*die Beschaffenheit, die der Prädicatvorstellung untersteht*,’ but of ‘*eine Beschaffenheit*,’ and in 1837, §.131 Bolzano unequivocally asserts that sometimes more than one property falls under the predicate concept in a truth. Nevertheless, Berg replies to the criticism of his singularity claim in Casari 1986, 61: ‘*In Bolzanos Wahrheitsdefinition ist [der Prädikatbegriff] eine Einzelvorstellung. Seine Ausführungen in §.131 lassen sich allerdings möglicherweise auch so auffassen, daß [der Prädikatbegriff] eine Gemeinvorstellung sein könnte*’ (BBGA 1.12/1, introduction of the editor, 12). It is hard to imagine a more devious concession to a critic. As a matter of fact, Bolzano’s *Ausführungen* leave no room whatsoever for any alternative *Auffassung*, and Berg’s statement about Bolzano’s definition of truth is simply false. And yet I believe that Berg is *substantially* right even if he is clearly wrong about what Bolzano says. (I hope it will soon become clear what I mean by this enigmatic statement.)

- (2) The sentence ‘A has b’ has no other sense than that every object falling under [the subject concept] A has one of the properties that fall under [the predicate concept] b;
- (3) and if there are several of the latter, it remains entirely undetermined which of them belongs to each A.<sup>85</sup>

Not everything is fine in this passage. In (2) Bolzano identifies inadvertently the sense of ‘A has b’ with the sense that the truth ascription ‘The proposition that A has b is true’ has according to (Df. T). In contrast to Frege, Bolzano maintains—and rightly so, I believe—that these sentences express *different* propositions.<sup>86</sup> An additional problem comes to the fore in the shaded segments of the text, and we shall confront it in the next paragraph but one. What matters with respect to the issue ‘Attribute Reading *vs.* Adherence Reading’ is contained in the unproblematic parts of (1) and (3) of the quotation.

If we are right in asserting that Socrates has courage then *exactly one* object falls under the subject concept of the proposition that is expressed, and our statement does indeed leave ‘entirely undecided’, as Bolzano puts it in (1), or ‘entirely undetermined’, as he puts it in (3), *which* of the various kinds of courage that (allegedly) fall under the predicate concept belongs to Socrates. But note that this would *not* be the case if the Adherence Reading were correct. The truth that Socrates has courage does *not* leave it open which of the many *adherences* that (according to the Adherence Reading) fall under the predicate concept of that truth adhere to the unique object that falls under its subject concept. Only one adherence is qualified for this role, and one can effortlessly read off from sentence ( $\Sigma$ ) which adherence it is: it is the courage of Socrates, not that of Jeanne, Rosa or Sophie. So the Adherence Reading is incompatible with Bolzano’s undecidedness claims.

Bolzano also tries to cover propositions that have more than one object falling under its subject concept. In this attempt he makes a mistake that surfaces in the shaded parts of (1) und (3). If Bolzano were right, then

---

85. (1) *Wir müssen ... sagen, daß es durch den Satz [dass A b hat] selbst ganz unentschieden bleibt, welche der mehren unter ihr [i.e. unter seiner Prädicatvorstellung] enthaltenen Beschaffenheiten (falls es derselben mehre gibt), dem einen oder den mehren der Vorstellung A unterstehenden Gegenständen zukomme. (2) Der Satz: A hat b, hat keinen anderen Sinn, als daß einem jeden der unter [dem Subjekt-Begriff] A stehenden Gegenstände eine der unter [dem Prädikat-Begriff] b begriffenen Beschaffenheiten zukomme); (3) läßt aber, falls es der letzteren mehre gibt, ganz unbestimmt, welche derselben es sey, die jedem A zukomme.* (1837, II 26f.; my numbering.)

86. Cp. Frege, e.g. 1918, 61 and Bolzano’s prophetic protest in 1837, I, 147 and 1851, § 13. I have discussed Frege’s redundancy thesis and various versions of truth-theoretical ‘nihilism’ in Künne 2003, Kap. 2.

an utterance of ‘Every pupil of Socrates has courage’ would only express a truth if there were a kind of courage such that *every* pupil of Socrates has *it*. But of course this proposition could also be true if there were *no* kind of courage such that every pupil of Socrates has courage of *that* kind.

Bolzano’s Caius example (C) shows that he reckons with the possibility that a property that (allegedly) falls under the predicate concept of a proposition is a common property of many actual objects. This supports the interpretive hypothesis that the properties that feature in Bolzano’s definition of truth are attributes even if the truth candidates are about actual objects. The word ‘has’ in the definiens of (Df. T) should have the same meaning, no matter whether we apply the definiens to propositions about one or all primes, or whether we apply it to propositions about one or all primates. Only one of the two meanings of ‘has’ I took care to distinguish is available in both types of case, namely *instantiates*. This concept is required for Bolzano’s theory of predication in general and of truth in particular. His definition of truth could be correct even if his ontological assumption that there are adherences were false. The concept *adhering*, on the other hand, is required for what he calls his *Metaphysik*, that is, for his ontology of actual objects.

At least one further worry remains with Bolzano’s definition of truth, —a worry that is independent of the obvious question whether (Df. T) is broad enough to cover all truth candidates.<sup>87</sup> Let me phrase it in terms of my favourite example. Which relation obtains between the concept of courage and the many kinds of courage? According to (Df. T) the latter ‘fall under’ the former. Is that correct?<sup>88</sup>

What falls under the concept *courage*, I dare say, is the property of being courageous—and nothing else. Of course, this uniqueness claim does not exclude that there are different kinds of courage. However, they do not fall under the concept *courage* but rather under the concept *a kind of courage*. Civil courage, for example, falls under this concept in the same way in which Socrates falls under the concept *a human being*. But civil courage does *not* fall under the concept *courage* anymore than the kind Indigenous Australian falls under the concept *a human being*. Therefore I propose to modify (Df. T):

---

87. That is a question that I have discussed (under the heading ‘the Procrustes Problem’) in my 2003, 111f.

88. In a long exchange of letters Edgar Morscher finally convinced me that this question should be answered negatively. This does not imply that he would approve of my argument for the negative answer.

(Df. T)\* A proposition is true if, and only if, every object that falls under its subject concept has **the** property that falls under its predicate concept.

If that is correct then Bolzano was wrong to deny in §.131 that the predicate concept is ‘taken in its entire extension (*nach ihrem ganzen Umfange genommen*)’ when one claims that A has b (1837, II, 26, 29). The entire (Bolzanian) extension of the concept *courage* is what you refer to when you give a true affirmative answer to the question, ‘What falls under it?’, and that is what you do when you reply, ‘The property of being courageous’. What remains of this section of *WL* is an important commentary on the revised definition of truth: if a property b is such that there are several kinds of b, then the proposition that (every) A has b is true just in case (every) A has some kind of b. The proposition leaves open *which* kind of b it is that makes the proposition true.

#### BIBLIOGRAPHY

- Adelung, Johann Christoph: *Grammatisch-kritisches Wörterbuch der Hochdeutschen Mundart*. Leipzig <sup>2</sup>1793–1801.
- Aquinas c.1255: *In Libros Sententiarum Magistri Petri Lombardi*. Parma: Fiacca-dori 1857.
- Aristoteles: *Topica, Categoriae, De Interpretatione, Analytica Posteriora, Metaphysica*. In: *Opera*. Immanuel Bekker (ed.), Berlin: Preußische Akademie der Wissen-schaften, 1831–1870.
- Baumgarten, Alexander Gottlieb 1766: *Metaphysik*. Transl. by Georg F. Meier. Halle (Saale): Hemmerde.
- Betti, Arianna 2012: “Bolzano’s Universe: Metaphysics, Logic and Truth”. In: Leila Haaparanta & Heikki Koskinen (eds.), *Categories of Being: Essays in Metaphysics and Logic*. Oxford: Oxford University Press, 169–190.
- Bolzano, Bernard: BBGA: *B. Bolzano Gesamtausgabe*. Jan Berg et al. (eds.), Stutt-gart: Frommann-Holzboog, 1969ff.
- 1805–09 (?): *Vorbereitungslehren*. In: BBGA, series 2A, vol. 5, Jan Berg (ed.), 1977, 77–132.
- 25. 02.1810, Sermon ‘*Uiber den Muth*’. In: BBGA, series 2A, vol. 17/1, Kurt F. Strasser (ed.), 1977, 248–76.

- 1810a: *Beyträge zu einer verbesserten Darstellung der Mathematik*, Prag: C. Widtmann.
- 1810b: *Allgemeine Mathesis*. In: BBGA, series 2A, vol. 5, Jan Berg (ed.), 1977, 13–64.
- 1833–41: *Einleitung und Erste Begriffe der allgemeinen Größenlehre*. In: BBGA, series 2A, vol. 7, Jan Berg (ed.), 1975.
- 1834: *Lehrbuch der Religionswissenschaft*. 4 Bde., Sulzbach: Seidel. In: BBGA, series 1, vols. 6/1–8/4, Jaromir Loužil (ed.), 1994–2006.
- 1837: *Wissenschaftslehre*. 4 vols., Sulzbach: Seidel. Reprint Aalen: Scientia <sup>2</sup>1981. In: BBGA, series 1, vols. 11/1–14/3, Jan Berg (ed.), (1994–2006). See also: Paul Rusnock & Rolf George.
- 1838: *Athanasia oder Gründe für die Unsterblichkeit der Seele* (1827). 2nd improved edn., Sulzbach: Seidel, Reprint Frankfurt/M: Minerva 1970.
- 1839: *Dr. Bolzano und seine Gegner*. Sulzbach: Seidel. Reprint Amsterdam: Rodopi 1970. In: BBGA, series 1, vol. 16/1, Jaromir Loužil (ed.), (1989).
- 1840–41: *Aphorismen zur Physik*. In: BBGA, series 2A, vol. 12/3, Jan Berg (ed.), 1978, 105–148.
- 1841: *Bolzano's Wissenschaftslehre und Religionswissenschaft in einer beurtheilenden Uebersicht*. Sulzbach: Seidel.
- 1851: *Paradoxien des Unendlichen*. Leipzig: Reclam. Christian Tapp (ed.), Hamburg: Meiner 2012.
- 1935: *Der Briefwechsel B. Bolzano's mit F. Exner*. Eduard Winter (ed.). Prag: Böhmisches Gesellschaft der Wissenschaften.
- 1980: *Philosophische Tagebücher, 1817–1827*. In: BBGA, series 2B, vol. 17, Jan Berg (ed.).
- 1981: *Philos. Tagebücher, 1811–1817*, 1. Teil. In: BBGA, series 2B, vol. 16/1, Jan Berg (ed.).
- 2005: *Briefe an František Přithonský 1824–1848*. In: BBGA, series 3, vol. 3/1–3, Jan Berg (ed.).
- 2006: *Briefe an J. Sommer u.a.*. In: BBGA, series 3, vol. 5/1, Jan Berg (ed.).
- 2009: *Philos. Tagebücher, 1803–1810*. In: BBGA, series 2B, vols. 14 & 15, Jan Berg (ed.).
- Casari, Ettore 1986: „Bemerkungen über die Bolzanosche Wissenschaftslehre“. In: *Logik und Grundlagenforschung, Festkolloquium zum 100. Geburtstag von Heinrich Scholz*. Münster: Schriftenreihe der Westfälischen Wilhelms-Universität, New Series, vol. 8, 53–66.
- 2006: „Some Remarks on Bolzano's Notion of a Quality“. In: Eduardo Ballo & Miriam Franchella (eds.), *Logic and Philosophy in Italy*. Monza: Polimetrica, 185–201.
- Descartes, René 1644: *Principia Philosophiae*. In: *Œuvres de Descartes*. Charles Adam & Paul Tannery (eds.), Paris 1897ff., Bd. VIII-1.

- Frege, Gottlob 1892: “Über Begriff und Gegenstand”. Reprinted (with original pagination) in his: *Funktion, Begriff, Bedeutung*. Günther Patzig (ed.), Göttingen: Vandenhoeck & Ruprecht, 71994.
- 1918: “Der Gedanke”. Reprinted (with original pagination) in his: *Logische Untersuchungen*. Günther Patzig (ed.), Göttingen: Vandenhoeck & Ruprecht, 52003.
- Goudin, Antoine 1671: *Philosophia Thomistica, iuxta inconcussa, tutissimaque Divi Thomae dogmata, Tomus Primus. Logica*. zit. nach der Ausgabe Madrid 1789.
- Kahn, Charles H. 1973: *The Verb BE in Ancient Greek*. Dordrecht: Reidel.
- Kant, Immanuel 1786: *Metaphysische Anfangsgründe der Naturwissenschaft*. [??]
- 1787: *Kritik der reinen Vernunft*, 2. Auflage. [??]
- Künne, Wolfgang 2003: *Conceptions of Truth*. Oxford: Clarendon Press.
- 2006: “Properties in Abundance”. In: Peter F. Strawson & Arindam Chakrabarti (eds.), *Universals, Concepts and Qualities*. Burlington: Ashgate, 249–300.
- 2007: *Abstrakte Gegenstände. Semantik und Ontologie* (1983), 2nd enlarged edn., Frankfurt/M.: Klostermann.
- Leibniz, Gottfried Wilhelm 1691: „*Animadversiones in partem generalem Principiorum Cartesianorum*“. In: *Die philosophischen Schriften von G. W. Leibniz*. Carl I. Gerhardt (ed.), reprint Hildesheim: Olms 1965, vol. 4, 354–406.
- 1704: *Nouveaux essais sur l’entendement humain*. In: *Leibniz Akademie-Ausgabe*, Reihe VI, Bd. 6.
- 1716: “5th Letter to Clarke”. In: Carl I. Gerhardt (ed.), vol. 7. Repr. Hildesheim: Olms 1965, 389–420.
- Morscher, Edgar 2004: „Kann es in der Welt 3 Indexikalität geben?“. In: Mark Siebel & Mark Textor (eds.), *Semantik und Ontologie*. Frankfurt/M: Ontos, 355–385.
- 2008: *Bernard Bolzano’s Life and Work*. Sankt Augustin: Academia.
- Ramsey, Frank P. 1925: “Universals”. Reprinted in D. H. Mellor & Alex Oliver (eds.), *Properties*. Oxford: Oxford University Press, 57–73.
- Reusch, Johann Peter: *Systema logicum antiquorum atque recentiorum item propria praecepta exhibens*. Jena 1734. Repr. Hildesheim 1990 (in: Christian Wolff, *Gesammelte Werke*, series III, vol. 26).
- Rusnock, Paul & George, Rolf 2014: *Bernard Bolzano, Theory of Science*. Translated by Paul Rusnock & Rolf George. Oxford: Oxford University Press.
- Schnieder, Benjamin 2002: *Substanz und Adhärenz—Bolzanos Ontologie des Wirklichen*. Sankt Augustin: Philosophia.
- 2004: *Substanzen und (ihre) Eigenschaften*. Berlin: De Gruyter.
- Simons, Peter 1987: *Parts. A Study in Ontology*. Oxford: Clarendon Press.
- 1997: “Bolzano on Collections”. In: Wolfgang Künne, Mark Siebel & Mark Textor (eds.), *Bolzano and Analytic Philosophy. Grazer Philosophische Studien* 53, 87–108.

- Simons, Peter 1999: „Bolzano über Wahrheit“. In: Edgar Morscher (ed.), *Bernard Bolzanos geistiges Erbe für das 21. Jahrhundert*. 13–27.
- 2006: “Austrian Philosophers on Truth”. In: Mark Textor (ed.), *The Austrian Contribution to Analytic Philosophy*. London: Routledge, 1–159.
- 2014: “An English Philosopher in Salzburg?”. In: Edgar Morscher & Peter Simons, *Joint Ventures in Philosophy*. Sankt Augustin: Philosophia, 159–69.
- Strawson, Peter F. 1990: “Two Conceptions of Philosophy”. In: Robert Barrett & Roger Gibson (eds.), *Perspectives on Quine*. Oxford: Blackwell, 310–318.
- Textor, Mark 1996: *Bolzanos Propositionalismus*. De Gruyter: Berlin.
- 2003: „Bolzanos Ontologie—Eine Übersicht“. In: Kurt F. Strasser (ed.), *Die Bedeutung Bernard Bolzanos für die Gegenwart*. Prag: Filosofia, 79–106.
- 2006: (ed.) *The Austrian Contribution to Analytic Philosophy*. London: Routledge.

## THE LOGIC OF TRUTH

Edgar MORSCHER  
University of Salzburg

### *Summary*

Georg Henrik von Wright developed what he called a “Truth-Logic” by adding a truth operator “T” and appropriate axioms and rules to a standard system of propositional logic. Much more attractive than this basic system of von Wright’s “Truth-Logic”, however, is its extension to quantificational contexts which allows highly interesting philosophical applications. The paper reveals several shortcomings of von Wright’s quantificational version of his “Truth-Logic” and shows how to avoid them without giving up the basic idea of a Logic of Truth.

### *Dedication*

During Peter’s time in Salzburg, we worked together on several projects. Some resulted in joint papers, for example those in our *Joint Ventures in Philosophy* (Academia Verlag: Sankt Augustin 2014). Many of our projects, however, have remained unpublished. One of these concerned Georg Henrik von Wright’s “Truth-Logic”, but only several pages (with incidental notes) have survived. So I would like to take this occasion to turn these notes into a complete essay and dedicate it to Peter.

### 1. *Introduction*

In volume III of his *Philosophical Papers*, entitled *Truth, Knowledge, and Modality*<sup>1</sup>, Georg Henrik von Wright develops a Logic of Truth or Truth-Logic, for short (26–41). It serves as the basis of von Wright’s so-called “Logic of Predication” (42–51) which consists in an extension of his Truth-Logic to predicate logic. So in fact von Wright’s Logic of Truth is a

---

1. Oxford 1984. Page numbers in the following text refer to this book.

Propositional Truth-Logic and his Logic of Predication is the corresponding Quantificational Truth-Logic. I will therefore abbreviate in what follows the former by 'PTL' and the latter by 'QTL'.<sup>2</sup>

PTL offers a stimulating and fruitful framework for important issues in philosophical logic. In section 2, I will sketch and evaluate the main ideas of PTL. In section 3, however, I will argue that QTL suffers from several short-comings and does not even satisfy von Wright's own aims and intuitions. In section 4, I will present a modification of von Wright's QTL and show that it is closer to his intentions.

## 2. Von Wright's PTL: An appraisal

Von Wright presents PTL as an axiomatic calculus. It differs from a standard system of propositional logic by including the symbol ' $T$ ' (which reads 'it is true that') in its otherwise standard vocabulary.

The vocabulary of PTL consists of sentential variables ( $p, q, r, \dots$ ), representing declarative sentences, two primitive sentential connectives ( $\neg$  for negation and  $\wedge$  for conjunction) and three additional sentential connectives ( $\vee$  for disjunction,  $\rightarrow$  for subjunction and  $\leftrightarrow$  for bijunction, defined in the usual way), the truth operator  $T$  and brackets. The definition of a well-formed formula of PTL results from adding the following clause to the usual definition: A well-formed formula of PTL, preceded by the letter ' $T$ ', is a well-formed formula of PTL.

Von Wright offers also a truth-table procedure vis à vis the formulas of PTL. Since in this "semantical" form (i.e., *via* truth-tables), PTL is easier to grasp than are its axioms and rules, I shall restrict myself to this variant of PTL. For this purpose three kinds of truth value are assigned

---

2. Concerning the relation between QTL and PTL, von Wright attached great importance to the following fact: Although in *Truth, Knowledge, and Modality*, the chapter on truth-logic (PTL) is followed by the chapter on the logic of predication (QTL), his paper "Logic of Predication", on which the corresponding chapter of the book is based, belongs to an earlier stage of his philosophical development than his paper "Truth and Logic". In a letter of 25 March 1992 he wrote me: "Von dem, was Sie zu meiner 'Logik der Prädikation' sagen, erkenne ich, dass ich, bei der Zusammenstellung des Bandes 'Truth, Knowledge, and Modality' eines ärgerlichen Versäumnisses schuldig bin. Ich hätte sagen sollen, dass der Aufsatz 'The Logic of Predication' zu einer viel früheren Phase meiner Entwicklung als Logiker gehört als der Aufsatz 'Truth and Logic'. Ich sehe den erstgenannten nicht als eine Art von 'Fortsetzung' des letzteren. (Beide haben allerdings einen gemeinsamen 'Urquell' in einem Aufsatz von mir aus den 50er Jahren über die Negation; er enthält manche Fehler, aber ist für mein späteres Denken fruchtbar gewesen.)"

to the sentential variables:  $t$  for ‘true’,  $f$  for ‘false’ and  $n$  for ‘neither true nor false’<sup>3</sup>.

The “meanings” of negation ( $\neg$ ) and truth ( $T$ ) are fixed by the following truth-tables:

TRUTH-TABLE I

$p$	$\neg p$	$Tp$
$t$	$f$	$t$
$n$	$n$	$f$
$f$	$n$	$f$

For conjunction ( $\wedge$ ) we have the following truth-table though von Wright does not state it explicitly:

TRUTH-TABLE II

$p \wedge q$	$t$	$n$	$f$
$t$	$t$	$n$	$f$
$n$	$n$	$n$	$f$
$f$	$f$	$f$	$f$

*Falsehood* of a proposition ( $F$ ) is defined, according to von Wright, as the truth of its negation (2), i.e.:

$$Fp =_{def} T\neg p$$

The non-primitive sentential connectives are defined as usual (26):

$$\begin{aligned}
 p \vee q &=_{def} \neg(\neg p \wedge \neg q) \\
 p \rightarrow q &=_{def} \neg(p \wedge \neg q) \\
 p \leftrightarrow q &=_{def} \neg(p \wedge \neg q) \wedge \neg(\neg p \wedge q)
 \end{aligned}$$

---

3. For practical reasons I prefer the small letters ‘t’, ‘f’, and ‘n’ to the numerals ‘1’, ‘0’ and ‘½’ used by von Wright for the three truth-values.

A well-formed formula of PTL is a *PTL-tautology* iff its truth-value is  $t$  for all assignments of truth-values to sentential variables occurring in it.

It can easily be verified by means of truth-tables that all axioms of von Wright's axiomatic version of PTL (as, e.g.,  $T\neg p \rightarrow \neg Tp$  and  $T(p \wedge q) \leftrightarrow (Tp \wedge Tq)$ ) are PTL-tautologies and that the following principle holds for all well-formed formulas  $s$  of PTL:

$s$  is a theorem of PTL iff  $s$  is a PTL-tautology, that is,  
 $\vdash_{\text{PTL}} s$  iff  $\models_{\text{PTL}} s$ .

Consider, now, the following simple truth-tables:

TRUTH-TABLE III

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$p$	$\neg p$	$Tp$	$\neg Tp$	$T\neg p$	$p \vee \neg p$	$T(p \vee \neg p)$	$Tp \vee T\neg p$	$T(p \vee \neg p) \leftrightarrow T(p \vee T\neg p)$	$Tp \vee \neg Tp$
$t$	$f$	$t$	$f$	$f$	$t$	$t$	$t$	$t$	$t$
$n$	$n$	$f$	$t$	$f$	$n$	$f$	$f$	$t$	$t$
$f$	$t$	$f$	$t$	$t$	$t$	$t$	$t$	$t$	$t$

TRUTH-TABLE IV

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$p$	$\neg p$	$Tp$	$\neg Tp$	$T\neg p$	$\neg(p \wedge \neg p)$	$T\neg(p \wedge \neg p)$	$\neg(Tp \wedge T\neg p)$	$\neg(Tp \wedge \neg Tp)$
$t$	$f$	$t$	$f$	$f$	$t$	$t$	$t$	$t$
$n$	$n$	$f$	$t$	$f$	$n$	$f$	$t$	$t$
$f$	$t$	$f$	$t$	$t$	$t$	$t$	$t$	$t$

These truth-tables yield a check on the correctness of von Wright's treatment of logical laws such as the Law of Excluded Middle, the Law of Bivalence, and the Law of Contradiction. By the Law of Excluded Middle von Wright understands the assertion that the disjunction of a proposition and its negation is true for every proposition, and by the Law of Bivalence he understands the assertion that any proposition is either true or false (33). However, the Law of Contradiction asserts, according to von Wright, that no proposition both is true and is false. These distinctions can be displayed by identifying the three logical laws mentioned with different formulas of PTL, namely:

Law of Excluded Middle (LEM):  $T(p \vee \neg p)$

Law of Bivalence (LB):  $Tp \vee Fp$ , i.e.,  $Tp \vee T\neg p$

Law of Contradiction (LC):  $\neg(Tp \wedge Fp)$ , i.e.,  $\neg(Tp \wedge T\neg p)$

While LC is a PTL-tautology (cf. truth-table IV, column 8), neither LEM nor LB is a PTL-tautology (cf. III 7 and III 8), but they are logically equivalent (cf. III 9):  $\models_{\text{PTL}} \text{LC}$ ,  $\not\models_{\text{PTL}} \text{LEM}$ ,  $\not\models_{\text{PTL}} \text{LB}$ ,  $\models_{\text{PTL}} \text{LEM} \leftrightarrow \text{LB}$ .<sup>4</sup>

These results coincide with von Wright's intuitions: According to him there are sentences which lack truth-value, nevertheless are meaningful and express propositions. There are different reasons why a declarative and meaningful sentence lacks a truth-value; it may lack truth-value because it is prescriptive or evaluative (22), or because it contains a category mistake (23), or because it is an open sentence containing a free variable (such as the sentence 'it is raining' or 'it is green'). Von Wright says the last express unspecified or generic propositions (24).

Von Wright's intuitions, however, demand a high price from PTL: None of the tautologies of standard propositional logic (PL) is a PTL-tautology because for at least one assignment of truth-values to the variables occurring in them they will all get the value  $n$  (cf. III 6 and IV 6). For von Wright this result of PTL shows that the laws of PL are not valid for all propositions (including those which lack truth-value), but only for propositions that are either true or false. The result itself, however, is strongly counter-intuitive. It could be repaired, at least in part, by weakening the definition of a PTL-tautology and requiring of a PTL-tautology only that it does not have the value  $f$  for any assignment of truth-values to the variables occurring in it (instead of requiring that it is true or has the value  $t$  for every assignment):

A well-formed formula of PTL is a *PTL-tautology\** iff for no assignment of truth-values to the sentential variables occurring in it its truth-value is  $f$ .

This suggestion, however, will result only in a partial repair: it saves all of the tautologies of PL and turns them into PTL-tautologies\*, but the sentences of PTL claiming that they are true will not follow and will remain contingent. They do not even become tautologies in the weaker sense of

---

4. This is the case because distribution for  $T$  over disjuncts holds in both directions:  $T(p \vee q) \leftrightarrow (Tp \vee Tq)$ .

PTL-tautologies\* because their truth-column will contain the value  $f$  at least once; this is a result of the operator's  $T$  turning the value  $n$  of the formula to which it is prefixed into  $f$ .

There are many other possible variations and improvements of PTL some of which are hinted at by von Wright himself. For example, one could weaken the definition of falsehood by identifying the falsehood of a proposition with the negation of its truth instead of identifying it with the truth of its negation as hitherto:

$$F^*p =_{def} \neg Tp$$

This would result in new variants of the logical laws such as:

$$\text{LB}^*: Tp \vee \neg Tp$$

$$\text{LC}^*: \neg(Tp \wedge \neg Tp)$$

both of which are PTL-tautologies (cf. III 10 and IV 9).

Even if these results are considered less than satisfactory, PTL certainly elucidates, or helps to elucidate our intuitions about basic logical laws and their interrelations. The inventor of PTL also uses it as a framework in which to examine other philosophical problems of impact such as

- (1) Aristotle's puzzle about tomorrow's sea battle and the problem of future contingency (1–13),
- (2) the problem of how to talk about propositions in general and propositions without truth-value in particular without mystifying them (14–25),
- (3) the problem of the redundancy of the phrase 'it is true that', by reflecting on the status of the formula  $T(Tp \leftrightarrow p)$  within PTL (32f.),
- (4) the problem of Dialectical Synthesis or the question whether and how we can make sense of the claim that there are propositions which are both true and false (36–39), and finally
- (5) the problem of antinomies such as the Liar, represented in PTL by the formula  $T(p \leftrightarrow Fp)$  or  $T(p \leftrightarrow T\neg p)$  (39–41).

In a paper entitled "Truth-Logics"<sup>5</sup>, published several years after *Truth, Knowledge, and Modality*, von Wright generalized his idea of a truth-logic

---

5. *Logique et Analyse* 30 (1987), 311–334.

by presenting a variety of logical systems which he called ‘truth-logics’ (now in the plural form). Now he distinguished four “truth-values”: UT (univocally true), UF (univocally false), TF (true and false) and NTNF (neither true nor false), which are defined as follows<sup>6</sup>:

$$\begin{array}{ll}
 UTp =_{def} Tp \wedge \neg T\neg p & \text{(univocal truth)} \\
 UFp =_{def} \neg Tp \wedge T\neg p & \text{(univocal falsehood)} \\
 TFp =_{def} Tp \wedge T\neg p & \text{(truth-value overlap)} \\
 NTNFp =_{def} \neg Tp \wedge \neg T\neg p & \text{(truth-value gap)}
 \end{array}$$

In a *classical* system neither truth-value gaps nor truth-value overlaps are admitted. A *paraconsistent* system admits truth-value overlaps, and a *paracomplete* system admits truth-value gaps. One can, therefore, distinguish four main types of truth-logics: (1) classical systems, (2) systems which are paraconsistent but not paracomplete; (3) systems which are paracomplete but not paraconsistent; and (4) systems which are paraconsistent as well as paracomplete.<sup>7</sup> This further development of von Wright’s truth-logic has no impact, however, on QTL that is the main concern in what follows.

### 3. Von Wright’s QTL: A Criticism

After having presented PTL, von Wright tries to extend his framework to quantificational contexts. Such an attempt would seem to be very promising. Unfortunately, von Wright’s attempt does not live up to its promise.

Von Wright starts by quoting Aristotle and stating that something may neither have a property (being white) nor its negation (being not-white). Instead of saying that something is not-white or has the complement of whiteness, one also might say it lacks the property of whiteness. This, however, is not the same as the assertion that it is not the case that the thing in question is white. To express this distinction formally, von Wright uses the operator  $T$  from PTL and distinguishes two kinds of negation of a sentence  $TWa$  (‘It is true, that  $a$  is white’), namely:  $\neg TWa$  and  $T\neg Wa$ , where the former formula is used to deny that  $a$  has whiteness but the latter formula is used to affirm that  $a$  lacks whiteness. Take  $a$  to be the number 7, then  $\neg TWa$  is true and  $T\neg Wa$  is false.

---

6. “Truth-Logics”, 313.

7. “Truth-Logics”, 314.

Von Wright now introduces the concept of the range of a property  $A$  and defines it as the set of all things which either have or lack this property (42, 45):

$$\text{Range}(A) =_{\text{def}} \{x \mid TAx \vee T\neg Ax\}$$

Von Wright's treatment of quantification is very unsystematic and fragmentary, and QTL, therefore, remains—as compared to PTL—a mere torso. One reason for this seems to be that von Wright is chary about quantifying into  $T$ -contexts and therefore restricts quantification from the beginning in an awkward way.

Despite of these shortcomings von Wright indicates how QTL could be used and applied in different ways quite fruitfully, as the following examples may illustrate:

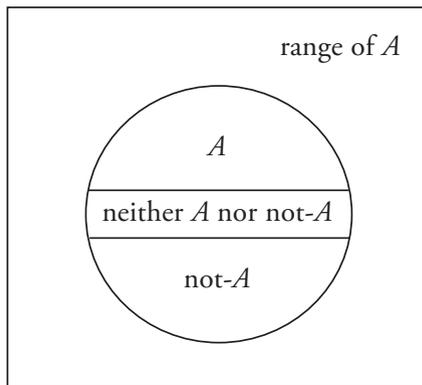
- (1) QTL can be used to explain what we mean when we say that everything has a certain property  $A$ . What we ordinarily mean by that is that there is something in the range of  $A$ , and everything in the range of  $A$  has  $A$ . Von Wright introduces a restricted universal quantifier in order to express this idea in QTL (46).
- (2) QTL can be used to clarify problems in connection with intuitionistic restrictions concerning the proofs for mathematical existence (47).
- (3) The treatment of fictional properties such as being a centaur can profit from QTL (48),
- (4) as well as the treatment of empty singular terms such as 'Pegasus' can (49f).
- (5) Von Wright finally (50f.) drops some hints vis à vis inference forms of QTL deviating from standard predicate calculus.

The main disadvantage of QTL is that it does not fully exist. Moreover, part of what exists does not even match von Wright's own intuitions. For example, consider von Wright's definition of the range of a property  $A$  as the set of things which either have or lack  $A$ . He himself presented and described examples which make clear that a thing can neither have nor lack a property  $A$  and still fall under the range of  $A$ , but this is excluded by his definition of range. Consider, e.g., the proposition that it is raining: "For some values of  $s$  and  $t$ ", von Wright states, "it is neither true nor false to say that it is raining in  $s$  at  $t$ " (37). Or consider contexts involving

the concepts of becoming, of change, and of process (37–39), or contexts of vagueness (39). Or, finally, consider colour properties where it often is natural to say of an object that it neither has nor lacks it (39) although it clearly belongs to the range of the property or of colour properties in general.

#### 4. *A modest Proposal to Save QTL*

Despite of this failure of QTL, I think it is easy to save its basic idea and purpose and to embed it into a semantics of predicate logic. In a standard semantics for predicate logic we usually start with a non-empty domain  $D$  and assign a subset of  $D$  to every one-place predicate. Applying a modified version of von Wright’s idea of the range of a property  $A$  (or of the corresponding one-place predicate) which matches his intuitions better than his own definition, we get the following picture:



This idea requires that the semantics has not just one assignment function  $\phi$  but three of them:  $\phi_1$ ,  $\phi_2$ , and  $\phi_3$ . This would result in the following proposal: We start with a system QL of quantificational logic whose vocabulary and well-formed formulas are defined as in standard systems and the semantics of which is modified in the following way (I restrict myself to one-place predicates):

An *interpretation of QL* is a quadruple  $\langle D, \phi_1, \phi_2, \phi_3 \rangle$  such that the following holds:

- (1)  $D \neq \emptyset$
- (2) for each one-place predicate  $\Phi$  of QL:  $\phi_1(\Phi) \subseteq D$  ( $\phi_1$  determines the range of  $\Phi$ )
- (3)  $\phi_2(\Phi) \subseteq \phi_1(\Phi)$  ( $\phi_2$  determines the set of things of which  $\Phi$  is true)
- (4)  $\phi_3(\Phi) \subseteq \phi_1(\Phi) \setminus \phi_2(\Phi)$  ( $\phi_3$  determines the set of things of which  $\Phi$  is false)
- (5) for each free individual symbol  $t$  of QL:  $\phi_2(t) \in D$

*Truth under an interpretation of QL:*

An atomic formula  $\Phi t$  of QL, consisting of a one-place predicate  $\Phi$  of QL and an individual symbol  $t$  of QL, is *true* under an interpretation  $\langle D, \phi_1, \phi_2, \phi_3 \rangle$  iff  $\phi_2(t) \in \phi_2(\Phi)$ ; and  $\Phi t$  is *false* under  $\langle D, \phi_1, \phi_2, \phi_3 \rangle$  iff  $\phi_2(t) \in \phi_3(\Phi)$ . The rest of the definition is as usual.

An atomic sentence  $\Phi t$  of QL can lack truth-value either because the object designated by the individual symbol ( $\phi_2(t)$ ) does not fall within the range of  $\Phi$  (i.e.,  $\phi_2(t) \notin \phi_1(\Phi)$ ) or because, although falling under its range, it does neither have nor lack the property in question, i.e.,  $\phi_2(t) \notin \phi_2(\Phi)$  and  $\phi_2(t) \notin \phi_3(\Phi)$ .

Now one can apply PTL within the meta-language to develop the semantics. If we add PTL to QL we get an improved version of QTL which we might call QTL\*. QTL\*, therefore, is the meta-linguistic framework by means of which we specify the truth-conditions for QL using PTL.

This line of thought, however, would result again—as does von Wright's own procedure—in ignoring a distinction which is important, given von Wright's examples and intuitions. Take an atomic sentence  $\Phi t$  consisting of a one-place predicate  $\Phi$  and an individual symbol  $t$  and consider the following two interpretations  $\langle D, \phi_1, \phi_2, \phi_3 \rangle$  and  $\langle D', \phi_1', \phi_2', \phi_3' \rangle$  such that:

- (i)  $\phi_2(t) \notin \phi_2(\Phi)$ ,  $\phi_2(t) \notin \phi_3(\Phi)$ , but  $\phi_2(t) \in \phi_1(\Phi)$
- (ii)  $\phi_2'(t) \notin \phi_2'(\Phi)$ ,  $\phi_2'(t) \notin \phi_3'(\Phi)$ , and also  $\phi_2'(t) \notin \phi_1'(\Phi)$

Under both interpretations  $\Phi t$  lacks truth-value, is neither true nor false. The important difference between the two cases, however, cannot be expressed within the framework developed so far. Nor will von Wright's system QTL suffice. Indeed, this might have been the reason that he dropped the third sub-class within the range of a predicate taking the

range as the union of the two sets of things of which the predicate is true (the set of things which have the corresponding property) or of which it is false (the set of things which lack the property).

The distinction can be saved by introducing a new semantic value. Consider first those sentences that are sensitive to truth-values even if eventually they turn out to be neither true nor false. This is the group of sentences which are either true (i.e., have value  $t$ ) or false (i.e., have value  $f$ ) or which are neither true nor false although being sensitive to truth-values (i.e., have value  $n$ ). From them we distinguish sentences that cannot have a truth-value to start with because they are insensitive to truth-values as, e.g., normative sentences or sentences containing category mistakes. For them I will introduce the semantic value  $i$  (i.e., insensitive to truth-values). Truth and truth-sensitivity under an interpretation of QL could then be defined as follows (I restrict myself again to atomic formulas):

An atomic well-formed formula  $\Phi t$  of QL is *truth-value sensitive* under an interpretation  $\langle D, \phi_1, \phi_2, \phi_3 \rangle$  iff  $\phi_2(t) \in \phi_1(\Phi)$ ;

$\Phi t$  is *true* under  $\langle D, \phi_1, \phi_2, \phi_3 \rangle$  iff  $\phi_2(t) \in \phi_2(\Phi)$ ;

$\Phi t$  is *false* under  $\langle D, \phi_1, \phi_2, \phi_3 \rangle$  iff  $\phi_2(t) \in \phi_3(\Phi)$ .

We may now distinguish the two cases mentioned earlier: in case (i),  $\Phi t$  is truth-valueless, but sensitive to truth-values (i.e., has value  $n$ ), in case (ii),  $\Phi t$  is insensitive to truth-values (i.e., has value  $i$ ). Representing this idea in the usual way would result in “truth-tables” of the following kind:

TRUTH-TABLE V

$p$	$\neg p$	$Tp$
$i$	$i$	$f$
$t$	$f$	$t$
$n$	$n$	$f$
$f$	$t$	$f$

TRUTH-TABLE VI

$p \wedge q$	$i$	$t$	$n$	$f$
$i$	$i$	$i$	$i$	$i$
$t$	$i$	$t$	$n$	$f$
$n$	$i$	$n$	$n$	$f$
$f$	$i$	$f$	$f$	$f$

Such a procedure, however, will “destroy” most of the nice results concerning logical laws established by von Wright’s Truth-Logic. Perhaps this was his motive for avoiding this additional step and stopping before these problems come out. My “outing” him shows that he corrupted his own intuitions, but that they can be saved without completely giving up the nice results concerning logical laws: We merely have to exempt interpretations under which sentences will get assigned the value  $i$  (i.e., sentences which are insensitive to truth-values) and restrict the results to sentences which are sensitive to truth-values.

The idea presented here to improve von Wright’s QTL can be generalized to extend to the treatment of arbitrary  $n$ -place predicates and in particular to nonhomogeneous predicates requiring different ranges for different places. A semantics taking care of these nonhomogeneous  $n$ -place predicates has to allow that each of its  $n$  places has its own range.

### *Appendix*

In a letter of 25 March 1992 to the author, Georg Henrik von Wright wrote me about my “modest proposal” and illustrated it with a nice example. Here is the text of his commentary:

Wie ich mich zu Ihrem ‘modest proposal’ stellen soll, habe ich (noch) nicht entscheiden können. Die Vierteilung des Domäns von Dingen in solche, die eine Eigenschaft  $A$  haben, in solche, die nicht- $A$  sind, solche, die weder  $A$  noch nicht- $A$  sind, und solche, die zu keiner von diesen drei Kategorien gehören, die aber doch – relativ zu  $A$  – ‘wahrheitswert-empfindlich’ sind, diese Vierteilung lässt sich gewiss machen und für weitere Entwicklungen auf interessante Weise benutzen, aber sie ist auch der weiteren Erklärung bedürftig.

Nehmen wir ein Beispiel:  $A$  soll die Eigenschaft 'rot' sein. Dinge, die wir ohne Zögern 'rot' nennen, haben die Eigenschaft  $A$ , Dinge, die eine bestimmte andere Farbe haben, sind nicht- $A$ . Welche sind nun die Dinge, die weder  $A$  noch nicht- $A$  sind? Der 'Ding-Begriff', so wie ich ihn in meinem Aufsatz benutze, soll 'alle Dinge' umfassen – also z.B. auch Zahlen oder Charaktereigenschaften, etc., etc. Also wäre auch die Zahl 5 ein Ding, das weder rot noch nicht-rot ist. Und in irgendeinem Sinne ist es wohl so. Aber dazu kann man sagen, dass eine derartige Klassifikation wichtige begriffliche Unterscheidungen verwischt. 'Interessanterweise' weder rot noch nicht-rot sind Dinge, die gefärbt sind, und zwar 'rötlich', von denen man aber doch nicht sagen möchte, sie seien 'rot', und auch nicht, sie seien 'nicht-rot', d.h. hätten eine *andere* Farbe.

Was ich im Aufsatz 'range' einer Eigenschaft nenne, könnte man auch so definieren, dass es (den 'Bereich') aller Dinge umfasst, die eine Eigenschaft  $D$  besitzen, die, wie der Begriff der Gefärbtheit, eine Dreiteilung in Dinge, die eine gewisse Farbe haben, die eine andere Farbe haben, und die sozusagen 'zwischen' die Farben fallen, zulässt.

Solch ein  $D$  fiele mit dem Zirkel in Ihrer Figur zusammen. Es ist also nicht das  $D$  des Vierecks. Wie soll man *dieses* auffassen? Nun kann man wie folgt weitergehen:

Manche Dinge sind grau. Sind sie 'farbig'? Und dieselbe Frage liesse sich auch für schwarze und weiße Dinge stellen. Solche Dinge reflektieren Licht und sind somit 'sehbar'. Aber 'farbig'? Man kann eine Entscheidung treffen und erklären, dass manche sehbare Dinge farbig, andere aber farblos sind. Sehbare Dinge wären dann ein Beispiel für *Ihr D* (das  $D$  im Viereck). Nun kann man noch einen Schritt machen:

Es gibt Dinge, die man zwar sehen kann, aber von denen es schwer ist zu sagen, ob sie als farbig oder nicht-farbig gelten sollen. Z.B. Dinge, die matt graubläulich sind. (Vorausgesetzt, dass graue Dinge nicht als farbig klassifiziert werden.)

Was möchte ich nun mit dem Obigen sagen? Ungefähr das Folgende: Die Einführung des Begriffs der 'truth-value sensitivity' ist die Einführung eines Begriffs eines übergeordneten *Genus* des Bereichs dessen, was Sie (nicht ich in meinem Aufsatz) 'range of a property' genannt haben. Das ist ein interessanter Schritt weiter auf dem Weg, den ich in meinem Aufsatz betreten hatte. Es gab eine Zeit, als ich mit dem Gedanken einer Erweiterung in diese Richtung umging – und in meinem Aufsatz über den Negationsbegriff aus den 50er Jahren habe ich skizzenhafte Andeutungen dazu gemacht. Ihr Aufsatz hat meine Gedanken wieder stimuliert. Aber ich fürchte, dass die Zeit für weitere Abenteuer auf diesem Gebiet für mich vorüber ist.

Mit herzlichem Gruss,  
Ihr Georg Henrik von Wright

## AN ANALYSIS OF LOGICAL DETERMINISM<sup>1</sup>

Jan WOLEŃSKI  
Jagiellonian University, Kraków

### *Summary*

Logical determinism is the view that some logical principles entail ontological determinism, that is, the view that the past uniquely forces the shaping of the future. The principle of bivalence and the law of excluded middle are usually considered crucial for logical determinism. On another interpretation, logical determinism is the view that the truth-value of a future contingent sentence is decided sempiternally. Jan Łukasiewicz's main motivation in inventing three-valued logic was to avoid logical determinism. The argument advanced in this paper intends to show that logical determinism is not derivable from classical logic.

### 1. *Introduction*

The problem addressed in this paper goes back to Aristotle and his considerations about tomorrow's sea battle. In a famous passage in (*De Interpretatione* 19a 25–30; in *The Works of Aristotle*, vol. 1: *Categoriae and De Interpretatione*, tr. By E. M. Edghill, Oxford University Press, Oxford 1928), the Stagirite says:

Everything must either be or not be, whether in the present or in the future, but it is not always possible to distinguish and state determinately which of these alternatives must necessarily come about.

Let me illustrate. A sea-fight must take place to-morrow or not, but it is not necessary that it either should not take place to-morrow, neither it is necessary that it should not take place, yet it is necessary that it either should or should not take place to-morrow. Since propositions correspond with facts, it is evident that when in future events there is a real alternative, and a

---

1. This paper is based on my talk delivered at the conference *Łukasiewicz in Dublin*, held at University College Dublin, July 1996. Peter Simons also participated in this meeting. Some ideas related to determinism and logic employed in this paper appears in (Woleński 2003; Woleński 2004; Woleński 2006).

potentiality in contrary directions, the corresponding affirmation and denial have the same character.

These words initiated a considerable and lengthy discussion (for historical information and substantial assessments, see Bernstein 1992; Cahn 1967; Gaskin 1995; Hintikka 1977; Jordan 1963; Karpenko 1990; Lucas 1989; Vuillemin 1996; Zagzebski 1991, Rice 2013). Basically, two problems were discussed. The first question was historical and concerned the way in which Aristotle should be interpreted. Did he revise logic or not? Literally understood, his text defends the principle of excluded middle but also tells us something about the logical value of future contingents, that is, sentences about accidental events which may or may not happen. The Stoics noted this problem as well. Their celebrated “Master argument” tried to prove that everything that is possible would happen in a more or less distant future. The second problem is systematic and concerns the question of how radical determinism or fatalism is related to some fundamental principles of logic, in particular, to the principle of bivalence. Roughly speaking, radical determinism (fatalism) is the view that the future is uniquely determined by the past or all future events are necessitated by the past.

The view that classical logic deductively implies radical determinism (**RD**) is called logical determinism. Jan Łukasiewicz himself did not use this label. More precisely, his view was (see Łukasiewicz 1922; Łukasiewicz 1970; see also Simons 1989 for a discussion of Meinong’s treatment of the excluded middle) that the law of the excluded middle (**EM**) or the principle of bivalence (**PB**), both interpreted metalogically, entail radical determinism when we add the principle of causality (**PC**).<sup>2</sup> Against (**RD**), Łukasiewicz constructed three-valued logic as a logical basis for indeterminism and free-will.<sup>3</sup> His reasoning goes like this: Take **PB + PC**. We have (the sign + replaces ‘and’ below)

(1) **PB + PC**  $\vdash$  **RD**.

On this account, rejecting **PB** or limiting **PC** amounts to canceling a sufficient condition for **RD**. But Łukasiewicz did more: He attacked both **PB** and **PC** as the sufficient condition because he introduced objective

---

2. See (Urchs 1992) for further analysis of this question.

3. Some qualifications are in order here: see (Jordan 1963). In particular, the early Łukasiewicz maintained that formal logic and ontology are very closely related. However, he abandoned this position later and denied that logic has any metaphysical consequences.

possibilities and thereby limited **PC**. It thus seems that a more satisfactory picture of the situation discussed by Łukasiewicz would be achieved by accepting the following principle

(2) **PB + PC**  $\dashv\vdash$  **RD**,

that is, by assuming that bivalence and the principle of causality are logically equivalent to radical determinism.

The term ‘logical determinism’ (more strictly its German counterpart: *logischer Determinismus*) was introduced by Schlick. Schlick spoke of the paradox of logical determinism as consisting in a conflict between logic and determinism:

[...] the principles of contradiction and excluded middle would not rank as statements about future states-of-affairs if determinism did not prevail. In fact, so Aristotle argued, if determinism is true, and if the future, therefore, is not already laid down and *determined* now, that it seems as if proposition ‘Event E will take place tomorrow’ could today be neither true nor false. (Schlick 1931, 202)

This problem was discussed by Waismann who, besides ‘logical determinism’, used the expressions ‘logical predestination’ or ‘logically preordained’ (Waismann 1956; Waismann 1965, 27–34; Waismann 1976, 59–68). To wit:

I shall single out for discussion [of the nature of philosophical problems—J. W.]—the question [...] whether the law of excluded middle, when it refers to the statements in the future tense forces us into a sort of logical Predestination. A typical argument is this. If it is true now that I shall do a certain thing tomorrow, say jump into the Thames, then no matter how fiercely I resist, strike out with hands and feet like a madman, when the day comes I cannot help jumping into the water; whereas, if this prediction is false now, then whatever efforts I may make, however many times I may nerve and brace myself, look down at the water and say to myself, ‘One, two, three’ – it is impossible for me to spring. Yet that the prediction is either true or false is itself a necessary truth, asserted by the law of excluded middle. From this the startling consequence seems to follow its is already now decided what I shall to do tomorrow, that indeed the entire future is somehow fixed, logically preordained. (Waismann 1956, 8f.)

There is a difference between Schlick and Waismann when it comes to singling out which logical principles are responsible for logical determin-

ism. Schlick speaks about the laws of contradiction and excluded middle, but Waismann only the latter. Waismann (1965; 1976) discussed the problem in the sections entitled “The Timeless Nature of Truth” and “Zeitlosigkeit der Wahrheit”. Nonetheless, neither Schlick nor Waismann accept logical determinism. Both reject the view that logic applies to the empirical world. Risking oversimplification, Schlick rejected this view because he thought of tautologies as devoid of content, whereas Waismann rejected it on the basis of a distinction between truth-bearers and truth-makers. In other words, both held that logic (Schlick, following one of the principal views of the Vienna Circle) and/or grammar (Waismann, influenced by Wittgenstein’s second philosophy) do not entail ontological consequences.

## 2. *Some basic intuitions*

Let us begin with **PB** and other related logical principles (see Woleński 2014 for a more comprehensive analysis). Following Łukasiewicz, one must distinguish very sharply the logical version of the laws of excluded middle and (non)-contradiction (**LC**) from the metalogical version. **PB** says that every sentence is either true or false. Clearly, it is a metalogical principle, and if we are to investigate its relation to **EM** and **LC**, both must also be formulated metalogically. We have ‘no sentence is jointly true and false, for **LC** and ‘every sentence is true or false’ for **EM**. From this we obtain:

$$(2) \text{ PB} \Leftrightarrow \text{LC} \wedge \text{EM}.$$

However, if we tacitly assume that we are working in a consistent system, **EM** suffices for arguments. And this in turn explains why this last principle is used as basic.

The next problem concerns truth and time. If truth is independent of time, then truth is absolute with respect to time (Twardowski 1902; I neglect other causes of the relativity of truth pointed out by Twardowski). This means that:

$$(3) \text{ if a sentence is true at a moment } t, \text{ it is also true at any other moment } t'.$$

Let us call this principle **AT**. **AT** can be divided into two parts:

- (3a) (eternality) if a sentence is true at  $t$ , it is also true at any  $t' \geq t$ ;  
 (3b)(sempiternality) if a sentence is true at  $t$ , it is also true at any  $t' \leq t$ .

In other words, the time-absoluteness of truth is the product of eternality and sempiternality.<sup>4</sup>

But what of sentences in the future tense, that is, what of future events, intuitively understood as contingent? Consider:

- (4) it will happen that  $A$  (e. g. ‘I will be in Warsaw on September 15, 2014’).<sup>5</sup>

I assume that (4) contains (or implies) as its essential factor:

- (5) it is possible that I will be in Warsaw on September 15, 2014.

Since we want to have (5) as expressing a contingent sentence, we add

- (6) it is possible that  $A$  and it is possible that not  $A$ .

Let the symbol  $\diamond A$  mean that  $A$  is contingent (= not necessary and not impossible). Thus, the contingency of  $A$  is defined by

- (7)  $\diamond A =_{df} \Diamond A \wedge \Diamond \neg A$  (or  $\neg \Box \neg A \wedge \neg \Box A$ ).

Accordingly, sentence (5) has the form  $\diamond A$  and is read as ‘it is contingent that I will be in Warsaw on September 15, 2014.

**AT** for temporal future sentences has the form:

- (8) a sentence of the form ‘it is possible that  $A(t')$ ’ is true at  $t$  iff it is true at any  $t' \neq t$

where  $t$  is a parameter related to the present time. Since  $t' \geq t$ , the symbol  $A(t')$  expresses that the sentence  $A$  concerns a future event.

If we thusly apply **AT** to (5), we get:

---

4. The term ‘sempiternal’ for the property of truth expressed in (3b) was coined by Peter Simons. The problem of the absoluteness of truth in Polish is informally discussed in (Woleński, Simons 1989).

5. In fact, I am writing this fragment on January 11, 2014.

- (9) (a) the sentence ‘it is possible that I will be in Warsaw on September 15, 2014’ is true today (= January 11, 2014) iff it is true at any  $t'' \leq t$  (sempiternality);
- (b) the sentence ‘it is possible that I will be in Warsaw on September 15, 2014’ is true today (= January 11, 2014) iff it is true at any  $t'' \geq t$  (eternality);
- (c) the sentence ‘it is possible that I will not be in Warsaw on September 15, 2014’ is true today (= January 11, 2014) iff it is true at any  $tt'' \leq t$  (sempiternality);
- (d) the sentence ‘it is possible that I will not be in Warsaw on September 15, 2014’ is true today (= January 11, 2014) iff it is true at any  $t'' \leq t$  (eternality).<sup>6</sup>

Clearly, the conditions listed in (9) are not at odds with the internal consistency of (9) modulo classical logic, although the conjunction ‘I will be in Warsaw on September 15, 2014 and I will not be in Warsaw on September 15, 2014’ is contradictory.

Having the basic intuitions specified, we can give the following explanations of logical determinism (**LD**):

- (10) (a) **LD** is derivable only by rules of logic;
- (b) **EM** implies **LD**;
- (c) **PB** implies **LD**;
- (d) **AT** implies **LD**;
- (e) it is logically impossible that there are any worlds with natural laws where the past is congruent with the actual world and yet have futures that diverges from the actual world;
- (f) it is logically necessary that worlds with the same past and the same natural laws have the same future;
- (g) the class of worlds with the same past is unique, that is, there are no different possible worlds;
- (h) the class of possible future worlds = {the future of the actual world};
- (i) logic excludes that ‘it is possible that  $A$  and it is possible that not- $A$ ’ is a factor of ‘it will be  $A$ ’, where  $A$  is neither tautological nor contradictory.

---

6. Leśniewski (1931) recommended avoiding the phrase ‘true at  $t$ ’. According to his view, we should include temporal coordinates into the contents of sentences. For instance, ‘ $a$  is  $b$  at  $t$  is true’ is preferable over ‘ $a$  is  $b$ ’ is true at  $t$ ’.

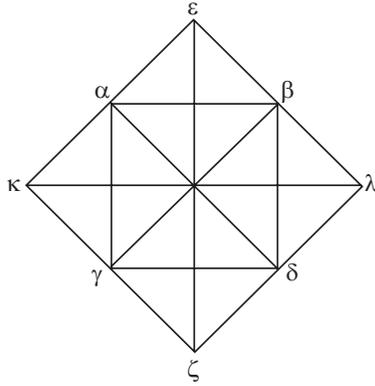
Without entering into the question whether (10a)–(10i) are exactly equivalent, let me assume that they are closely related.<sup>7</sup> In what follows, I will argue that none of the versions of **LD** mentioned are plausible inasmuch as none of them are derivable from logic itself. This does not mean that I opt for logical indeterminism understood as a negation of **LD** in one of its formulations. My position is that the issue of determinism/indeterminism is independent of logic. As a result, the truth of future contingencies as defined in (9) is consistent with determinism as well as indeterminism. Just like the young Łukasiewicz, I think that logic is ontologically applicable. In other words, logic can be understood as an applied ontology. But I will argue, *pace* Schlick and Waismann, that logic itself does not entail any ontological consequences. This holds whether or not one regards tautologies as formulas devoid of factual content. While I agree with Łukasiewicz that **RD** = **PB** + **PC**, I don't think this situation should force us to change logic.

### 3. *A modal analysis of determinism*

Let me start with an octagonal diagram ( $\Delta$ ) of oppositions, which generalizes a well-known traditional logical square used for displaying formal relations holding between the categorical sentences ('every *S* is *P*', 'some *S* are *P*', etc.) or modalized sentences 'it is necessary that *A*', 'it is possible that *A*' and their negated forms (see Woleński 1996; Woleński 2008 for the construction of this diagram and its various philosophical applications; for instance, analogical relations hold for deontic modalities):

---

7. It is easily seen that not all explanations of **LD** are mutually equivalent. For example, (10b) is stronger than (10c), because **EM** is a part of **PB**. On the other hand, the role of **LC** in **PB** consists in excluding a trivial derivation (inconsistency entails everything). Thus, **EM** is more important for the discussed problem than **LC**.



$\Delta$  is interpreted as follows:

$\alpha$   $A$  is determined ( $\mathbf{D}A$ );

$\beta$   $\neg A$  is determined ( $\mathbf{D}\neg A$ );

$\gamma$   $\neg(\neg A$  is determined) ( $\neg\mathbf{D}\neg A$ );

$\delta$   $\neg(A$  is determined) ( $\neg\mathbf{D}A$ );

$\varepsilon$   $A$  is determined or  $\neg A$  is determined ( $\alpha \vee \beta$ ;  $\mathbf{D}A \vee \mathbf{D}\neg A$ );

$\zeta$   $\neg(\neg A$  is determined)  $\wedge$   $\neg(A$  is determined) ( $\gamma \wedge \delta$ ;  $\neg\mathbf{D}\neg A \wedge \neg\mathbf{D}A$ ;  $A$  is contingent);

$\kappa$   $A$  (it is true that  $A$ );

$\lambda$   $\neg A$  (it is true that  $\neg A$ ).

We have the following logical dependencies (they constitute the minimal core of modal rules, structurally similar to some theorems of first-order logic):

- (11) (a)  $\alpha \Rightarrow \varepsilon$ ;  
 (b)  $\alpha \Rightarrow \gamma$ ;  
 (c)  $\beta \Rightarrow \varepsilon$ ;  
 (d)  $\beta \Rightarrow \delta$ ;  
 (e)  $\neg(\alpha \wedge \beta)$ ;  
 (f)  $\gamma \vee \delta$ ;  
 (g)  $\neg(\alpha \Leftrightarrow \delta)$ ;  
 (h)  $\neg(\beta \Leftrightarrow \gamma)$ ;  
 (i)  $\zeta \Rightarrow \gamma$ ;  
 (j)  $\zeta \Rightarrow \delta$ ;

- (k)  $\neg(\varepsilon \leftrightarrow \zeta)$ ;
- (l)  $\alpha \vee \beta \vee \zeta$ ;
- (m)  $\varepsilon \vee \zeta$ ;
- (n)  $\kappa \Rightarrow \gamma$ ;
- (o)  $\lambda \Rightarrow \delta$ ;
- (p)  $\kappa \vee \lambda (A \vee \neg A)$ ;
- (q)  $\alpha \Rightarrow \kappa$  (if  $A$  is determined,  $A$  is true);
- (r)  $\beta \Rightarrow \lambda$  (if  $\neg A$  is determined,  $\neg A$  is true).

It is important that the converses of (q) and (r) do not hold. Also the formulas  $\alpha \vee \beta$  and  $\gamma \wedge \delta$  do not express logical principles. We also see that the operator related to  $\gamma$  is not homogenous, because it can be interpreted as  $\alpha \vee \kappa \vee \zeta$ , intuitively ‘ $A$  is determined’ (necessary in a sense) or true or contingent’. From (11) we obtain:

- (12) (a)  $\mathbf{D}A \Rightarrow \Diamond A$ ;  $\Box A \Rightarrow \Diamond A$ ;
- (b)  $\mathbf{D}A \Rightarrow \mathbf{T}A$ ;  $\Box A \Rightarrow \mathbf{T}A$ ;
- (c)  $\mathbf{T}A \Rightarrow \Diamond A$ ;
- (d)  $\blacklozenge A \Rightarrow \Diamond A$ .

The converses of (12a)–(12b) do not hold on purely logical grounds, at least, if  $\Delta$  displays our logic.

Furthermore, we define (recall that  $A$  is neither tautological nor contradictory):

- (13) (a) *radical determinism* (**RD**):  $\forall A(\varepsilon), \forall A(\alpha \vee \beta), \forall A(\mathbf{D}A \vee \mathbf{D}\neg A)$ ;
- (b) *radical indeterminism* (**RI**):  $\forall A(\gamma \wedge \delta), \forall A(\neg \mathbf{D}A \vee \neg \mathbf{D}\neg A)$ ;
- (c) *moderate determinism* (**MD**):  $\exists A(\alpha \vee \beta) \wedge \exists A(\gamma \wedge \delta)$ ;  
 $\exists A(\mathbf{D}A \vee \mathbf{D}\neg A) \wedge \exists A(\neg \mathbf{D}A \vee \neg \mathbf{D}\neg A)$ ;
- (d) *moderate indeterminism* (**MI**):  $\exists((\gamma \wedge \delta) \wedge \exists A(\alpha \vee \beta))$ ;  
 $\exists A(\neg \mathbf{D}A \wedge \neg \mathbf{D}\neg A) \wedge \exists A(\mathbf{D}A \vee \mathbf{D}\neg A)$ ;
- (e) *minimal determinism* (**DM**):  $\exists A(\alpha), \exists A \neg \mathbf{D}A$ ;
- (f) *minimal indeterminism* (**IM**):  $\exists A(\zeta), \exists A(\neg \mathbf{D}A \wedge \neg \mathbf{D}\neg A)$ ;
- (g) *determinism with respect to a fact  $f$*  (**D<sup>f</sup>**):  $\mathbf{D}f \vee \mathbf{D}\neg f$ ;
- (h) *indeterminism with respect to a fact  $f$*  (**I<sup>f</sup>**):  $\neg \mathbf{D}f \wedge \neg \mathbf{D}\neg f$ .

Intuitively, **RD** says that everything is determined, **RI** that everything is not determined, **MD** that something is determined and something else

is not determined, **MI** that something is not determined and something else is determined, **DM** that at least something is determined, **IM** that at least something is not determined, **D<sup>f</sup>** that an individually specified fact *f* is determined, and **I<sup>f</sup>** that an individually fact *f* is not determined. **MD** and **MI** are logically not distinguishable, but can differ in the distribution of what is determined and what is not (hence, I introduced another succession in the respective formulas, but this has no logical relevance). **D<sup>f</sup>** entails **DM** and **I<sup>f</sup>** entails **IM**, **DM** follows from **RD** and **MD**, **IM** follows from **RI** and **MI**.

None of the statement in (13a)–(13h) expresses a logical truth.<sup>8</sup> This gives a preliminary argument for the thesis that logic and determinism are independent. It can be strengthened by an observation coming from (12). **RD** suggests that **DA** and  $\diamond A$  are equivalent. On the other hand, **RI** motivates that  $\diamond A$  and  $\blacklozenge A$  are equivalent (modifications for the rest of the possible positions are immediate.) Clearly, these equivalences exceed logic. The concept of possibility is in a special position, because  $\diamond A$  is entailed by  $\Box A$ , **TA** and  $\blacklozenge A$ . Hence, we conclude that the existence of possible events is consistent with **RD**, **RI**, **MD**, **MI**, **DM**, **IM**, **D<sup>f</sup>** and **I<sup>f</sup>**, that is, with all the possible kinds of determinism and indeterminism listed in (13).

#### 4. *Logic and determinism*

**RD** is traditionally expressed the sentence ‘everything is necessary’. This statement can be formalized as (I will use the box  $\Box$  instead the letter **D**; recall that necessity and other modalities can be interpreted ontologically)

$$(14) \quad \forall A(\Box A \vee \Box \neg A).$$

To prove **LD** by the devices of logic as the only means of inference from a theorem of logic as the only premise, the box should be interpreted as logical necessity. The form of (14) suggests an appeal to **EM** as the starting point of an argument. Thus, assume this principle, that is, the formula

---

8. In general, I make no specific assumptions about determinism, necessity and related concepts. My arguments work for any mechanism invented for showing that the future is dependent on the past as well as for an absolute or relative understanding of necessity. Additional constraints will be employed on special occasions, for example, when necessity should be interpreted as logical.

$A \vee \neg A$ . Apply necessitation to derive (\*)  $\Box(A \vee \neg A)$ . However, (\*) does not imply (14). This fact, very often invoked (for instance, see Wessel 1999, 151f.), does not depend on logical or metalogical understanding of **EM**.<sup>9</sup> Thus, deriving **LD** from (14) is a simple logical mistake, because the latter is not a logical rule. In particular, it does not express **EM**.<sup>10</sup>

Although the above reasoning is elegant in its simplicity it still does not solve the question. The argument starts in pure logic and shows that something fails to be derived from it. However, we rather need to show that **LD** does not work in logic or metalogic interpreted ontologically. To that end, we should interpret **EM** in terms of  $\Delta$ . Let the symbol  $\blacktriangle$  refer to one of  $\Box$ ,  $\Diamond$ , **T** or  $\blacklozenge$ . Schematically speaking, **EM** has the form

$$(15) \quad \forall A(\blacktriangle A \vee \neg \blacktriangle A).$$

with its special instances obtained by replacing the symbol  $\blacktriangle$  by one of the members of the set  $\{\Box, \Diamond, \mathbf{T}, \blacklozenge\}$  (15) has a clear ontological import, contrary to the purely logical form, i. e., the formula  $A \vee \neg A$ . In order to justify **LD**, we should prove that **RD**, that is  $\forall A(\Box A \vee \Box \neg A)$ , follows from at least one instance of (15). We can skip the formula  $\forall A(\Diamond A \vee \neg \Diamond A)$ , because it is consistent with all positions and thereby cannot serve as the only basis for **RD**. Also, the formula  $\forall A(\blacklozenge A \vee \neg \blacklozenge A)$  does not work, because the statement that everything is necessary cannot follow from the sentence that everything is contingent or is not contingent. Eliminating  $\forall$  we obtain  $\blacklozenge A \vee \neg \blacklozenge A$ . The only route left is to apply the proof by cases, but it is blocked by the first disjunct. The same concerns the formula  $\forall A(\Box A \vee \neg \Box A)$ , because we need the equivalence  $\neg \Box A \vee \Box \neg A$ , which begs the question, because it equates non-necessity with impossibility.

To be fair, (15) instantiated by  $\Box$ ,  $\Diamond$  or  $\blacklozenge$  was never used for proving **LD** or questioning it, unless **EM** was rejected for other reasons. Thus, there remains

$$(16) \quad \forall A(\mathbf{T}A \vee \neg \mathbf{T}A),$$

---

9. The converse entailment holds. The steps from (13) to (\*) and from the latter to **EM** are justified by modal logic.

10. This mistake is made in Taylor (1962; this paper tries to prove fatalism from logic; (Rice 2013) uses the label 'logical fatalism'). See (Garson 2013, 153–156) for a detailed analysis of Taylor's error.

but we must devote some attention to an interpretation of **T** in the present context before using this form of **EM**. Of course, we cannot understand truth as an epistemological category. Let us read **TA** as ‘it is real that *A*’; I will use the symbol **RA** for this interpretation and change (16) into

$$(17) \quad \forall A(\mathbf{R}A \vee \mathbf{R}\neg A).$$

Since (12b) and (12c) are fairly plausible for the reading recommended by (17), we can accept that necessity entails reality and that reality entails possibility. The operator **R** behaves differently than  $\Box$ ,  $\Diamond$  and  $\blacklozenge$ , because it admits the equivalence of  $\neg\mathbf{R}A$  and  $\mathbf{R}\neg A$ , which allows us to replace (17) by

$$(18) \quad \forall A(\mathbf{R}A \vee \mathbf{R}\neg A)$$

Intuitively, the formula  $(\bullet)\neg\mathbf{R}A \Leftrightarrow \mathbf{R}\neg A$  expresses that not being real is equivalent to being unreal just as ‘not true’ means ‘false’. In fact,  $(\bullet)$  can be transformed into the formula

$$(19) \quad \forall A(\mathbf{R}A \vee \mathbf{R}\neg A) \wedge \neg(\mathbf{R}A \wedge \mathbf{R}\neg A).$$

This last formula can be interpreted as the ontological principle of bivalence (**OPB**). It says that everything is real or unreal and nothing is jointly real and unreal. Now we clearly see that neither (18) nor (19) entail **RD** without an additional premise

$$(20) \quad \forall A(\mathbf{R}A \Rightarrow \Box A),$$

that is, without equating reality and necessity. Of course, we can derive various things from  $\forall A(\Box A \vee \Box\neg A)$ , in particular,  $\forall A\Box(A \vee \neg A)$  (see below),  $\forall A\Diamond(A \vee \neg A)$  and  $\forall A\mathbf{R}(A \vee \neg A)$ , but not the formula  $\forall A\blacklozenge(A \vee \neg A)$ . In fact, if everything is determined, necessities are necessary, possible and real (necessity implies reality and possibility), but not contingent. Yet these observations give no support for **LD**.

To complete the issue of **EM** and its relation to **LD**, we should investigate whether admitting contingencies as occurring ontological items introduces any troubles. Assume that something is contingent, that is

$$(21) \quad \exists A\blacklozenge A \Leftrightarrow \exists A(\neg\Box\neg A \wedge \neg\Box A).$$

By eliminating  $\exists$  from the first formula in brackets, we obtain

$$(22) \quad \neg \Box \neg f \wedge \neg \Box f.$$

Define  $\neg \Box \neg f$  as  $\Diamond f$ . This gives

$$(23) \quad \Diamond f \wedge \Diamond \neg f.$$

Since  $\Diamond A \wedge \Diamond B$  does not entail  $\Diamond(A \wedge B)$ , we cannot conclude that contradictions are possible, which would threaten **EM** and **PB**. This observation closes the issue of how classical logic is related to determinism. The answer is: in no way. Otherwise speaking, any direct derivation of determinism (the same concerns indeterminism) from pure logic consists in conflating the latter with ontology.

By contrast, let us investigate now what happens when **PB** is limited. This move, according to Łukasiewicz, opens the door for indeterminism. If exceptions to **PB** are admitted, for example, some sentences, in particular, future contingents, are valued by the third value **n**, (17) and (18) are not equivalent and become contingent.<sup>11</sup> Let us work with (17). Consequently, its negation  $(\bullet) \neg \forall A(\mathbf{R}A \vee \mathbf{R}\neg A)$  is contingent as well. Simple syntactic transformations give  $(\bullet\bullet) \exists A(\neg \mathbf{R}A \wedge \neg \mathbf{R}\neg A)$  and  $(\bullet\bullet\bullet) \neg \mathbf{R}f \wedge \neg \mathbf{R}\neg f$ , which contradict classical logic, because  $\mathbf{R}A$  and  $\neg \mathbf{R}\neg f$  are classically equivalent. To save the situation one must change logic. Let  $f$  be a future contingency. Thus, both  $f$  and  $\neg f$  are possible and not necessary as well as unreal. Since Łukasiewicz's three-valued logic validates the formula  $\Diamond A \wedge \Diamond B \Leftrightarrow \Diamond(A \wedge B)$ , we get that  $\Diamond(f \wedge \neg f)$ . This means that *some* contradictions are possible. Formally, there is no contradiction, because if  $f$  and  $\neg f$  have the value **n**, their conjunction  $f \wedge \neg f$  has the same status. Hence, the formula  $\Diamond(f \wedge \neg f)$  can be true. However, some authors (see Prior 1953) complain about this consequence by arguing that contradictions are always false. It seems that Łukasiewicz had a special reason for his solution, namely, that he wanted to express contingencies directly in the future tense without using possibility as an auxiliary device. On the other hand, there is a special problem with the concept of reality in ontology based on Łukasiewicz' logic. For some reason, he rejected necessary objects, at least in the empirical world, but this was a minor point. Yet

---

11. See (Karpenko 1990) for a comprehensive study of the problem of determinism in the context of many-valued logics.

we can ask whether Łukasiewicz's framework admits of contingent past events. Assume that  $\mathbf{R}f \wedge \diamond f \wedge \diamond \neg f$ . This entails  $\mathbf{R}f \wedge \diamond(f \wedge \neg f)$ . However, this result seems strange when applied to the past, unless we extend the third logical value also to sentences about past contingencies.

### 5. *Contingencies and classical logic*

Until now, I considered 'being real' as a homogenous predicate. However, something more should be said in order to achieve its full interpretation in the language of classical logic. Due to (12b), the realm of necessity falls under being real, but the former does not exhaust the latter, at least from a purely conceptual point of view. ( $\Delta$ ) does not suggest anything about the relation between the real and the contingent. In fact, reality is ambiguous, because it can be determined or contingent. Let us look once more at my visit in Warsaw on September 15, 2014. According to the assumed analysis, it is future and contingent at every time before my arrival to Warsaw. On the other hand, this statement must somehow be supplemented by a qualification of my visit after its realization. Is it real and still contingent? It seems that the following description is adequate (I use the predicative form instead employing operators):

- (24)  $A$  is real and contingent ( $\mathbf{R}^c A$ ) iff  $A$  is real,  $\neg A$  is not real,  $A$  was possible and  $\neg A$  was possible.

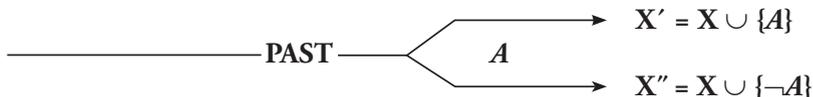
Assume that  $\mathbf{R}^c A$  holds for some  $A$ . Hence,  $A$  cannot be necessary and thereby  $\Box A$  is false. This suffices for the truth of the implication  $\Box A \Rightarrow \mathbf{R}A$  and shows that it is universally valid. The best interpretation of (24) consists in applying to past contingent events. Thus, the real grows through history.

To accommodate contingencies into classical logic without any problems, I will identify possible world with models of maximally consistent sets and use the concept of branchability (Asser 1972, 168f.) as defined by:

- (25) (a) a set  $\mathbf{X}$  of formulas branches at a formula  $A$  iff the sets  $\mathbf{X} \cup \{A\}$  and  $\mathbf{X} \cup \{\neg A\}$  are consistent;  
 (b) a set  $\mathbf{X}$  is branchable iff there is a formula  $A$  at which  $\mathbf{X}$  branches;

- (c) a set  $\mathbf{X}$  is branchable iff  $\mathbf{X}$  is a consistent and incomplete set of sentences.

Let  $\mathbf{X}$  be a consistent set of sentences and  $A$  be a formula (sentence) independent of  $\mathbf{X}$ . Thus, due to the independence of  $A$  with respect to  $\mathbf{X}$ , this set is incomplete and the sets  $\mathbf{X}' = \mathbf{X} \cup \{A\}$  and  $\mathbf{X}'' = \mathbf{X} \cup \{\neg A\}$  are consistent. Consequently, the conditions listed in (25) are fulfilled and  $\mathbf{X}$  branches at  $A$ . Note that we do not need to assume that the sets  $\mathbf{X}'$  and  $\mathbf{X}''$  are maximally consistent. We can display this situation by the diagram ( $\Delta'$ )



Now we shall adopt an ontological interpretation.<sup>12</sup> Let  $A$  be a sentence uttered at  $t$  fixed as present and refers to a future event, and let  $\mathbf{X}$  consist of all true sentences about the past. Since  $\mathbf{X}$  contains only true sentences, it is consistent. It is plausible to assume that  $A$  is independent of  $\mathbf{X}$ . If  $A$  is independent,  $\mathbf{X}'$  and  $\mathbf{X}''$  are consistent and have models, let us say,  $\mathbf{W}'$  and  $\mathbf{W}''$ , and let **PAST** be the model of  $\mathbf{X}$ . We can change the diagram ( $\Delta'$ ) into ( $\Delta''$ )



**PAST** can be considered as the initial segment of  $\mathbf{W}'$  and  $\mathbf{W}''$ . Otherwise speaking,  $\mathbf{W}'$  and  $\mathbf{W}''$  enlarge **PAST**; intuitively **PAST** comprises everything that happened until the moment  $t$  (including this moment itself). Moreover,  $\mathbf{W}'$  and  $\mathbf{W}''$  are parts of different possible worlds, that is, models of consistent sets. Let us call them full ontological models. They are different, because one, with  $\mathbf{W}'$  as its segment validates  $A$ , but the second, related to  $\mathbf{W}''$  verifies  $\neg A$ . **PAST** is, of course, the initial segment of the real world  $\mathbf{W}^R$ , which grows through time. This concurs with an earlier observation (see (24)) that the real contingent is growing. Depending on what will happen, **PAST** will enlarge,

12. I follow Woleński (2012).

at first, to  $\mathbf{W}'$  or  $\mathbf{W}''$ , but then to other worlds, relatively to further branchings.<sup>13</sup>

Now we are ready to state the following truth-condition for future contingencies. The standard condition for sentences preceded by the operator  $\diamond$  says that  $\diamond A$  is true,  $A$  is true at least in one possible world accessible from the world identified as the real one. Consequently, the sentence  $\blacklozenge A$  is true, if there two possible worlds  $\mathbf{W}'$  and  $\mathbf{W}''$ , accessible from the real world, such that  $A$  is true in one of those worlds, but  $\neg A$  in the other. It is enough to fix **PAST** as our real world (or its part).

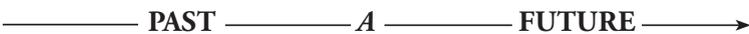
- (26) a sentence ‘it is contingent at  $t \leq t'$  that  $A(t')$ ’ is true in  $\mathbf{W}^{\mathbf{R}}$  (= **PAST**) at  $t$  iff there are two  $\mathbf{W}'$  and  $\mathbf{W}''$  enlargements of **PAST** such that  $A$  is true in  $\mathbf{W}'$  iff  $\neg A$  is true in  $\mathbf{W}''$ .

We can also consider possible past scenarios. In particular, (26) can be easily transformed for sentences of the form  $\mathbf{R}^{\mathbf{C}}A$  in the following way:

- (27)  $\mathbf{R}^{\mathbf{C}}A(t')$  is true in **PAST** in  $t \geq t'$  iff  $A$  is true in **PAST**,  $\neg A$  is false in **PAST**, but there is a possible world  $\mathbf{W}'$  such that  $A(t')$  is true at  $t'$ .

Yet real contingencies can be reduced to future contingencies, because if a sentence is a real contingent, then it means that it had to be a future contingency at some moment  $t$  in the past.

Must we acknowledge that **PAST** is branchable? We need not necessarily make this proviso. Instead, we can adopt the simplest model, illustrated by the diagram ( $\Delta'''$ )



This last ontological model displays the main point of **RD**: there is only one course of events. Thus, **RD** is equivalent to the assertion that the **PAST** is not branchable, but branchability automatically introduces a restriction of radical determinism. Finally, radical indeterminism could perhaps be illustrated by a model in which reality consists of separate and unconnected points. Now **LD** is the view that we can decide on purely logical grounds that ( $\Delta'''$  offers us a correct description of the world. However, this is fairly

13. Needless to say, we consider the relation of **PAST** to a selected future contingent event. This implies nothing about the question of how the future exists: in advance or not, in full detail or not, etc.

impossible, similar to proving indeterminism (or moderate determinism) by logic. The decision between particular ontological models depends on assumptions concerning the scope and character of causality and eventually as well as other aspects of regularity, for instance, statistical distributions or fractals. Hence, all models, like  $(\Delta')$  or  $(\Delta''')$  are conditional and assert nothing about the actual (meta)physical structure of  $\mathbf{W}^R$ . Speaking otherwise, logic offers various possible (formal) models for ontology, but it cannot decide which one is realized in the world.

Leśniewski (1912) argued that (3a) and (3b) are logically equivalent; namely, a sentence  $A$  is eternally true if and only if it is sempiternally true. Leśniewski's original proof is fairly complicated, but it can be very simply reproduced in our framework; I will deal with future contingencies only. Assume that  $A$  is true at a moment  $t$ . At first, we will prove that sempiternality entails eternity. If  $A$  is true sempiternally, it is true at every  $t \leq t'$ . Since the branching moment is critical, we assume that  $t$  is just this parameter. Consequently,  $A$  is true in  $\mathbf{W}'$  or  $A$  is true in  $\mathbf{W}''$ . Furthermore,  $A$  is true in  $\mathbf{M}' = \mathbf{PAST} + \mathbf{W}'$  or  $A$  is true in  $\mathbf{M}'' = \mathbf{PAST} + \mathbf{W}''$ . This implies that  $\neg A$  is false in  $\mathbf{M}' = \mathbf{PAST} + \mathbf{W}'$  or  $\neg A$  is false in  $\mathbf{M}'' = \mathbf{PAST} + \mathbf{W}''$ . Clearly,  $A$  cannot change its logical value in a model belonging to  $\{\mathbf{M}', \mathbf{M}''\}$  without producing inconsistency. Thus, if  $A$  is sempiternally true, it is eternally true as well. To prove the converse implication, suppose that  $A$  is eternally true, that is, if  $A$  is true at  $t$ , it is also true at any moment  $t' \geq t$ ; of course,  $A$  is true in  $\mathbf{W}'$  or  $\mathbf{W}''$  and *a fortiori*, in  $\mathbf{M}'$  or  $\mathbf{M}''$ . Consider a moment  $t'' \leq t$ . Assume that  $A$  is false at  $t''$ . This means that  $A$  is false in **PAST**. However, this implies that  $A$  is false in  $\mathbf{M}'$  or  $\mathbf{M}''$ , contrary to our earlier assumption. Thus, if  $A$  is eternally true, it is sempiternally true as well. This result is interesting. Many philosophers, Łukasiewicz is a good example in this respect, accept the eternity of truth as uncontroversial, but complain about sempiternality. The argument for the logical equivalence of (3a) and (3b) shows that something is wrong in this view, because if sempiternality is felt as non-intuitive, the same should concern eternity. It seems that here we have another instance of a conflation of logic and metaphysics. In general, the equivalence of (3a) and (3b) demonstrated that contingent truth is absolute.<sup>14</sup> Since we are working with the concept

---

14. Note, however, that Łukasiewicz's logic does not force abandoning the idea that truth is absolute. He could argue that although the third value becomes truth or falsity depending of what will happen, truth (or falsehood) never becomes the third value (or truth). Clearly, the absoluteness of truth (and falsehood) is restricted to its eternity. Moreover, this position introduces instability of logical values in general.

of truth in a model, we can consider the semantic notion of truth as absolute (Placek 2006). Leibniz proposed the principle *unumquodque quando est, oportet est* (what is, is necessary). However, assuming that the concept of truth satisfies (3) does not determine that it is necessary or contingent. Thus, we should rather adopt a weaker principle, namely *unumquodque quando est, reale est* (what is, is real). Otherwise speaking, what happened could be necessary or contingent. In particular, (6) does not entail that every possibility sooner or later will be realized. This disproves the Master Argument.

#### REFERENCES

- Asser, Günther 1972: *Einführung in die mathematische Logik, Teil II, Prädikatenkalkül der ersten Stufe*. Leipzig: Teubner.
- Bernstein, Mark 1992: *Fatalism*. Lincoln: University of Nebraska Press.
- Cahn, Steven 1967: *Fate, Logic and Time*. New Haven: Yale University Press.
- Gaskin, Richard 1995: *The Sea Battle Argument and the Master Argument. Aristotle and Diodorus Cronos*. Berlin: de Gruyter.
- Garson, James 2013: *What Logics Mean. From Proof Theory to Model-Theoretic Semantics*. Cambridge: Cambridge University Press.
- Hintikka, Jaakko (in collaboration with Unto Remes & Simo Knuuttila) 1977: *Aristotle on Modality and Determinism*. Amsterdam: North-Holland.
- Jordan, Zbigniew 1963: "Logical Determinism". *Notre Dame Journal of Formal Logic* 4, 1–38.
- Karpenko, Alexander 1990: *Fatalism i slučajnost' buducevo. Logičeskij analiz* (Fatalism and Contingency of the Future. A Logical Analysis. In Russian). Moskva: Nauka.
- Leśniewski, Stanisław 1912: "Czy prawda jest tylko wieczna, czy też wieczna i odwieczna?" (Is All Truth only True Eternally or It Is True without a Beginning. In Polish), *Nowe Tory* 18, 493–512. English translation in: Stanisław Leśniewski 1992: *Collected Works*. Dordrecht: Kluwer, 86–114.
- 1931: "O podstawach matematyki". (On the Foundations of Mathematics. In Polish) §§10–11, *Przegląd Filozoficzny* 34, 142–170. English translation in: Stanisław Leśniewski 1992: *Collected Works*. Dordrecht: Kluwer, 351–382.
- 1992: *Collected Works*. Dordrecht: Kluwer.

- Lucas, John 1989: *The Future. An Essay on God, Temporality and Truth*. Oxford: Basil Blackwell.
- Łukasiewicz, Jan 1922: "On Determinism". In: Jan Łukasiewicz 1970: *Selected Works*. Amsterdam: North-Holland, 110–128. (English version of Łukasiewicz's Rector Speech in the academic year 1922/1923).
- 1930: "Philosophische Bemerkungen zu mehrwertigen Systemen des Aussagensenkalkül". *Comptes Rendus des Séances de la Société et de Lettres de Varsovie, Classe III* 23, 51–77. English translation in: Jan Łukasiewicz 1970: *Selected Works*. Amsterdam: North-Holland, 153–178.
- 1970: *Selected Works*. Amsterdam: North-Holland.
- Placek, Tomasz 2006: "A Puzzle about Semantic Determinism: Łukasiewicz's "On Determinism"—Thirty Years Later". In: Jacek Jadacki & Jacek Pasniczek (eds.), *The Lvov-Warsaw School – The New Generation*. Amsterdam: Rodopi, 171–185.
- Prior, Arthur 1953: "Three-Valued Logic and Future Contingents". *The Philosophical Quarterly* 3, 317–326.
- Rice, Hugh 2013: „Fatalism“. In: Edward Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), URL = <<http://plato.stanford.edu/archives/spr2013/entries/fatalism/>>.
- Schlick, Moritz. 1931: "Die Kausalität in den gegenwärtigen Physik", *Naturwissenschaften* 19, 145-162. English translation in: Moritz Schlick: *Collected Papers*, vol. II. Dordrecht: Reidel, 176-209.
- Simons, Peter 1986: "Łukasiewicz, Meinong, and Many-Valued Logic". In: Klemens Szaniawski (ed.), *The Lvov-Warsaw School and the Vienna Circle*. Dordrecht: Kluwer, 249–291. Reprinted in: Peter Simons 1993: *Philosophy and Logic in Central Europe from Bolzano to Tarski. Selected Essays*. Dordrecht: Kluwer, 193-226.
- Taylor, Richard 1962: "Fatalism". *Philosophical Review* 71, 56–66.
- Twardowski, Kazimierz 1902: "Über sogenannte relative Wahrheiten". *Archiv für systematisch Philosophie* 8, 415-447. English translation in: Kazimierz Twardowski 1999: *On Actions, Products and Other Topics in Philosophy*. Amsterdam: Rodopi, 147–169. Urchs, Max 1992: "On Determinism—Still Unproved in Classical Logic". *Bulletin of the Section of Logic* 21, 168-170.
- Vuillemin, Jules 1996: *Necessity or Contingency. The Master Argument*. Stanford: CSLI Publications.
- Waismann, Friedrich 1956: "How I See Philosophy?". In: Hywel Lewis (ed.), *Contemporary British Philosophy*. III. H. D. Lewis, George Allen and Unwin, London, 345–380. Reprinted in Friedrich Waismann 1968: *How I See Philosophy*. London: Macmillan, 1–38.
- 1965: *The Principles of Linguistic Philosophy*. London: Macmillan.
- 1976: *Logik, Sprache, Philosophie*. Stuttgart: Reclam. (This book was written in 1939).

- Wessel, Horst 1999: *Logik und Philosophie*. Berlin: *λογος*.
- Woleński, Jan 1996: "Logical Squares: Generalizations and Interpretations". In: Timothy Childers, Petr Kolár & Ondrej Majer (eds.), *Logica '95. Proceedings of the 9<sup>th</sup> International Symposium*. Praha: Filosofia, 67–75.
- 2003: "Dieterminism i logika". (Determinism and Logic. In Russian). *Voprosy Filosofii* 5, 71–81.
- 2004: "Zwei Beiträge zur logischen Philosophie". In: Bende Christiansen & Uwe Scheffler (eds.), *Was folgt. Themem zu Wessel*. Berlin: *λογος*, 299–309.
- 2006: "Three Contributions to Logical Philosophy". In: Jacek Jadacki & Jacek Pasniczek (eds.), *The Lvov-Warsaw School—The New Generation*. Amsterdam: Rodopi, 195–213.
- 2008: "Applications of Squares of Oppositions and Their Generalizations in Logical Analysis". *Logica Universalis* 2(1), 13–29. Reprinted in: Jan Woleński 2011: *Essays on Logic and Its Applications in Philosophy*. Frankfurt am Main: Peter Lang, 255–269.
- 2012: "Truth Is Eternal if and only if It Is Sempiternal". In: Edwin Tegtmeier (ed.), *Studies in the Philosophy of Herbert Hochberg*. Frankfurt: Ontos, 223–230.
- 2014: "An Abstract Approach to Bivalence". *Logic and Logical Philosophy* 23(1), 3–14.
- Woleński, Jan & Simons, Peter 1989: "De Veritate: Austro-Polish Contributions to the Theory of Truth from Brentano to Tarski". In: Klemens Szaniawski (ed.), *The Lvov-Warsaw School and the Vienna Circle*. Dordrecht: Kluwer, 391–453.
- Zagzebski, Linda 1991: *The Dilemma of Freedom and Foreknowledge*. Oxford: Oxford University Press.