

Was zählt eine Publikation?

Möglichkeiten zur Durchführung einer Forschungsevaluierung im Bereich der Betriebswirtschaftslehre

Gerhard Reichmann

Universität Graz
Institut für Informationswissenschaft und Wirtschaftsinformatik
Universitätsstraße 15/F 3, A-8010 Graz
E-Mail: gerhard.reichmann@uni-graz.at

Zusammenfassung

Im Rahmen des gegenständlichen Beitrages, der an den von Wolf Stock im Jahr 2000 verfassten Beitrag mit dem Titel „Was ist eine Publikation?“ anknüpft, wird diskutiert, wie im Rahmen einer Forschungsevaluierung für den Bereich der Betriebswirtschaftslehre vorgegangen werden könnte. Zunächst werden Einzelpersonen, also einzelne Professoren, die an einschlägigen Fakultäten tätig sind, als optimale Untersuchungseinheiten festgelegt. Als optimale Quelle zur Erhebung der benötigten Publikationsdaten erweisen sich von den Forschern selbst erstellte Publikationslisten, da diese am ehesten vollständig sind. Während man im Falle einer möglichst umfassenden Forschungsevaluierung sämtliche Einträge in diesen Publikationslisten als Basis für die Erstellung von Ranglisten heranziehen könnte, bietet sich bei Konzentration auf das zentrale Publikationsmedium (im Bereich der betriebswirtschaftlichen Forschung) eine Beschränkung auf Zeitschriftenbeiträge an. Die hier im Mittelpunkt stehende Frage „Was zählt eine Publikation?“, d.h. mit welchem Gewicht soll nun jede einzelne der für die Evaluierung relevanten Publikationen in die Evaluierung eingehen, kann sehr unterschiedlich beantwortet werden. Das vorgestellte Beispiel eines auf vier Personen reduzierten Forschungsrankings reicht bereits aus, um die Sensitivität der Forschungsevaluierung in Abhängigkeit von der Art der Gewichtung der relevanten Publikationen zu verdeutlichen. Empfohlen wird jedenfalls eine Berücksichtigung von Koautorenschaften sowie eine Gewichtung nach Zeitschriften, sofern entsprechende Zeitschriftenrankings vorhanden sind. Interessant ist, dass Stock in seiner beinahe 20 Jahre alten Publikation bereits nahezu alle im Rahmen dieses Beitrages diskutierten Punkte behandelt hat, wodurch sich seine Publikation als weitgehend zeitlos erweist.

Schlüsselwörter: Betriebswirtschaftslehre, Forschungsevaluierung, Forschungsranklisten, Gewichtung von Publikationen, Zeitschriftenrankings.

1 Einleitung

Im Universitätsbereich hat sich die Evaluierung¹ als Instrument zur Leistungsmessung mittlerweile weitgehend etabliert. Entsprechend den drei Kernbereichen universitärer Tätigkeit fokussiert die Evaluierung dabei entweder auf die Forschungs-, die Lehr- oder die Verwaltungstätigkeit. Auch im deutschsprachigen Raum hat es dahingehend in den vergangenen Jahren eine Vielzahl von Evaluierungen gegeben (zur Evaluierung der Verwaltung: vgl. z.B. Reichmann/Sommersguter-Reichmann 2010; zur Evaluierung der Lehre: vgl. z.B. Reichmann 2007). Besonders intensiv wurde im Bereich der Forschung evaluiert (vgl. dazu beispielsweise Albers 2015, Fülbi/Weller 2011 oder auch Wolf et al. 2005), besonders großen Bekanntheitsgrad genießt das „Handelsblatt-BWL-Ranking“². Die dabei ermittelten Leistungen, im Kern handelt es sich hierbei um den Forschungsoutput in Form von Publikationen, wurden und werden nicht nur als Nachweis für erfolgreiche Forschungstätigkeit herangezogen, sondern zunehmend auch als Basis für Leistungsvergleiche zwischen den evaluierten Einheiten. Universitäten, die bei derartigen Rankings gut abschneiden, haben in der Folge oftmals einen deutlichen Wettbewerbsvorteil, da sie bei der Vergabe von Budgetmitteln bevorzugt behandelt werden bzw. von Wissenschaftlern als attraktive (potentielle) Arbeitgeber wahrgenommen werden. Bezieht sich ein solches Forschungsranking auf die kleinste mögliche Einheit, also die einzelnen Forscher, so kann dieses für deren weitere Karriere von großer Bedeutung sein.

Aus evaluierungstheoretischer Sicht (vgl. dazu beispielsweise Wottawa/Thierau 2003 oder Stockmann/Meyer 2014) handelt es sich bei Forschungsevaluierungen in der Regel um summative, d.h. im Nachhinein durchgeführte Evaluierungen, die sich meist auf einen eindeutig begrenzten Zeitraum beziehen und vorwiegend im Hochschulbereich eingesetzt werden. Objekt der Evaluierung sind – zumindest auf Ebene der Datenerhebung – einzelne natürliche Personen, die Forschung betreiben. Bei Forschungsevaluierungen ist besonders darauf zu achten, dass diese unvoreingenommen durchgeführt werden. Günstig erscheint in diesem Zusammenhang die Form der Fremdevaluierung durch Außenstehende sowie jene der offenen Evaluierung, bei der die Ergebnisse publiziert werden. Dies spricht wiederum dafür, dass die Außenstehenden selbst Forscher sind, da diese üblicherweise auch im Hinblick auf Forschungsevaluierungen ein großes Interesse daran haben, diese Forschungsergebnisse zu publizieren. Zudem legen Forscher gewöhnlich großen Wert auf Reliabilität, wodurch die Transparenz der Ergebnisse gewährleistet wird.

Zentraler Anknüpfungspunkt von Forschungsevaluierungen ist die Forschungsleistung einzelner Wissenschaftler. Diese manifestiert sich üblicherweise in Form von Publikationen, die wiederum von anderen Forschern zitiert werden, sofern sie für deren Forschung von entsprechender Relevanz sind. Im Rahmen einer Forschungsevaluierung stellt sich zunächst die grundsätzliche

¹ Bei den beiden Begriffen „Evaluierung“ und „Evaluation“ handelt es sich um Synonyme. Während ersterer früher in Österreich vorherrschend war, in den letzten Jahren allerdings langsam von zweitem verdrängt wurde, fand in Deutschland praktisch ausnahmslos der Begriff „Evaluation“ Verwendung.

² Vgl. <https://www.handelsblatt.com/politik/konjunktur/bwl-ranking/> (27.2.2019).

Frage, was eine Publikation im Sinne dieser Evaluierung ist bzw. sein soll. Zählt jeglicher Eintrag in Publikationslisten bzw. Forschungsdatenbanken als Publikation oder sollten weitere (qualitative) Kriterien berücksichtigt werden? Stock (2000) ging dieser Frage bereits vor fast 20 Jahren nach und hat sie sehr umfassend und weitgehend disziplinenübergreifend beantwortet. Er ist dabei bereits sehr detailliert auf die Frage eingegangen, was eine einzelne, für eine Forschungsevaluierung als relevant erachtete Publikation zählen soll., die im Mittelpunkt der gegenständlichen Ausführungen steht. Es geht dabei v.a. um die Frage, ob und wie unterschiedliche, als relevant identifizierte Publikationen im Rahmen einer Forschungsevaluierung gewichtet werden sollen. Dabei wird untersucht, inwieweit die von Stock behandelten Aspekte aus heutiger Sicht noch vollständig bzw. aktuell sind.

Im zweiten Abschnitt werden zunächst mögliche Untersuchungsobjekte einer Forschungsevaluierung diskutiert, bevor im dritten Abschnitt auf die Problematik der Auswahl geeigneter Datenquellen zur Ermittlung des Forschungsoutputs einzelner Forscher eingegangen wird. Anschließend wird untersucht, welche Publikationen für eine Forschungsevaluierung im Bereich der Betriebswirtschaftslehre relevant sind. Im zentralen fünften Abschnitt wird diskutiert, was eine einzelne relevante Publikation zählen kann bzw. soll. Insbesondere werden dabei die Aspekte Koautorenschaft, Umfang einer Publikation sowie Zeitschriftenrankings behandelt. Die Sensitivität der Evaluierungsergebnisse in Abhängigkeit von den in den Abschnitten drei bis fünf vorgestellten Vorgangsweisen wird anhand eines Beispiels demonstriert. Zu diesem Zweck wird für eine kleine Gruppe von Forschern, bei denen es sich um die vier derzeit am Institut für Informationswissenschaft und Wirtschaftsinformatik der Universität Graz – dieses ist im Bereich der Betriebswirtschaftslehre angesiedelt – tätigen Professoren handelt (diese werden in der Folge als Forscher A, B, C und D bezeichnet), eine Forschungsevaluierung auf Basis der Publikationen durchgeführt. Das Beispiel, welches sich auf einen zehnjährigen Beobachtungszeitraum (2008-2017) bezieht, wurde aus ressourcenökonomischen Überlegungen auf die kleinstmögliche Stichprobenanzahl reduziert, die gerade noch ausreicht, um die Auswirkungen unterschiedlicher Vorgangsweisen zu verdeutlichen. Diese Untersuchung ist zudem eine Pilotstudie für die geplante Forschungsevaluierung für den Bereich der Betriebswirtschaftslehre in Österreich, die alle (ca. 300) in diesem Bereich tätigen Professoren, also auch die vier in der Pilotstudie betrachteten Personen, umfassen soll.³ Den Abschluss der Ausführungen bildet ein knappes Resümee.

2 Untersuchungsobjekt einer Forschungsevaluierung

Untersuchungseinheiten einer Forschungsevaluierung im Hochschulbereich können einzelne Forscher, aber auch Forschergruppen (z.B. Arbeitsgruppen oder Abteilungen eines Instituts), Institute bzw. Departments, Fakultäten bzw.

³ Im Gegensatz zu den oben ausgesprochenen Empfehlungen handelt es sich also sowohl bei der hier präsentierten Pilotstudie als auch bei der geplanten Hauptstudie um keine Fremdevaluierung durch Außenstehende, da die Forschungsleistung des Autors dieses Beitrages ebenfalls Gegenstand der Betrachtungen ist. Dies hat den Vorteil, dass der Autor mit dem untersuchten Bereich entsprechend vertraut ist.

Fachbereiche, gesamte Universitäten (bzw. sonstige Hochschulen) oder sogar Gruppen von Universitäten (vgl. dazu etwa Morgan 2004) sein. Die für die Evaluierung notwendigen Daten, also die Forschungsleistungen, werden üblicherweise auf Ebene der einzelnen Forscher erhoben und gegebenenfalls (wenn es sich bei den Untersuchungseinheiten eben nicht um einzelne Forscher handelt) auf Ebene der Untersuchungseinheiten aggregiert, da es in der Praxis kaum möglich ist, z.B. die gesamten Publikationen eines Instituts, einer Fakultät oder einer Universität in einem Schritt zu erheben, da die notwendigen Daten meist nicht in einer derartig aggregierten Form abrufbar sind.

Üblicherweise werden im Rahmen von Forschungsevaluierungen nicht nur einzelne Untersuchungseinheiten isoliert betrachtet, sondern mehrere Untersuchungseinheiten miteinander verglichen, was häufig zur Erstellung sogenannter Forschungsrankings führt. Die Gesamtheit der miteinander verglichenen Untersuchungseinheiten, z.B. sämtliche Forscher einer Universität, sämtliche Institute einer Fakultät, sämtliche Universitätsinstitute eines Fachbereichs in einem Land, alle (vergleichbaren) Universitäten in einem Land oder sogar – wie bei internationalen Universitätsrankings durchaus üblich – alle (vergleichbaren) Universitäten der Welt, bilden zusammen mit den im Mittelpunkt der Untersuchung stehenden Forschungsleistungen, z.B. Publikationen, Zitationen (als Beispiele für Forschungsrankings auf Basis von Zitationen: vgl. Dilger/Müller 2012 oder auch Dyckhoff et al. 2005), Auszeichnungen bzw. (Nobel)Preise und/oder Mitgliedschaften in Editorial Boards (vgl. dazu etwa Rost/Frey 2011), immer öfter auch die Akquise von Drittmitteln, das Untersuchungsobjekt der jeweiligen Forschungsevaluierung.

Die nachfolgenden Ausführungen und auch das gewählte Beispiel sind auf Forschungsleistungen in Form von Publikationen beschränkt. Zudem wird speziell auf den Forschungsbereich der Betriebswirtschaftslehre eingegangen. Im Beispiel werden die Forschungsleistungen von (vier) Professoren eines Instituts einander gegenübergestellt, d.h. Untersuchungseinheiten sind die einzelnen Forscher, Untersuchungsobjekt ist die Forschungsleistung in Form von Publikationen am Institut für Informationswissenschaft und Wirtschaftsinformatik der Universität Graz im Betrachtungszeitraum (2008-2017). Als Untersuchungsziel kann die Erstellung verschiedener (in Abhängigkeit von der gewählten Vorgangsweise) Forschungsrankings für dieses Institut genannt werden.

Die oben angesprochene geplante Forschungsevaluierung für den gesamten Bereich der Betriebswirtschaftslehre in Österreich soll sich ebenfalls auf Publikationen und einen Betrachtungszeitraum von 10 Jahren beschränken. Untersuchungseinheiten könnten hierbei nicht nur die einzelnen Forscher, sondern auf höher aggregierten Stufen auch die einzelnen Institute oder die einzelnen Universitäten (beschränkt auf den Fachbereich der Betriebswirtschaftslehre) sein. Im Falle eines Vergleiches auf Instituts- oder Universitätsebene stellt sich die Frage, ob es fair und sinnvoll ist, kleinere und größere Institute bzw. Universitäten undifferenziert miteinander zu vergleichen. Ein Institut, dem 5 Professoren angehören, müsste in so einem Fall deutlich besser dastehen als ein Institut, an dem nur 2 Professoren tätig sind, da auch im Forschungsbereich durchaus economies of scale zu beobachten sind (vgl. z.B. Johnes 1995). Im Rahmen von internationalen Forschungsrankings wird diesbezüglich erstaunlicherweise oft undifferenziert vorgegangen. Vermutlich deshalb, weil eine sachge-

rechte Differenzierung nur schwer möglich ist und jedenfalls hohen Ressourceneinsatz erfordert. Die naheliegende Lösung, die Forschungsleistung eines Instituts oder einer Universität einfach durch die Anzahl der dort tätigen Professoren zu dividieren, führt bei genauerer Betrachtung nicht unbedingt zu einer Verbesserung der Vorgangsweise, da es an den verschiedenen Instituten bzw. Universitäten auch eine vollkommen unterschiedliche Anzahl an wissenschaftlichen und nichtwissenschaftlichen Mitarbeitern geben kann, was wiederum großen Einfluss auf die Forschungsleistungen der Untersuchungseinheiten haben kann. Wenn man das obige Beispiel weiterdenkt, so könnte es sein, dass es am Institut, dem die 5 Professoren angehören, nur 5 sonstige wissenschaftliche (z.B. Assistenten) und 2 nichtwissenschaftliche Mitarbeiter (z.B. Sekretäre) gibt, während am Institut, dem die 2 Professoren angehören, noch 20 wissenschaftliche (etwa zum Teil über Drittmittelprojekte finanziert) und 5 nichtwissenschaftliche (z.B. auch Bibliothekare, die Recherchen für die Forschung durchführen können, und EDV-Spezialisten, die für die Forschung notwendige Software bedienen können) tätig sind. Im Gegensatz zu einfachen Verhältniskennzahlen in Form der oben erwähnten Teilproduktivitäten bieten aktuellere Benchmarking-Methoden, wie z.B. die Stochastic Frontier Analysis oder die Data Envelopment Analysis (vgl. z.B. Bogetoft/Otto 2011), durchaus Möglichkeiten, derartige Unterschiede adäquat in die Untersuchung einfließen zu lassen.

3 Datenquellen einer Forschungsevaluierung

Untersuchungen haben gezeigt, dass die Wahl der Datenquelle für die Ergebnisse einer Forschungsevaluierung von größter Bedeutung ist (vgl. Meyer et al. 2012). Möchte man eine solche auf Basis von Publikationen für den Bereich der Betriebswirtschaftslehre durchführen, so bieten sich in erster Linie vier mögliche Datenquellen an (vgl. dazu auch Dorsch et al. 2018):

- a) die Publikationslisten der Forscher,
- b) die Datenbanken des Web of Science,
- c) die Datenbank "Scopus" sowie
- d) die Suchmaschine „Google Scholar“.

Wirklich empfehlenswert ist eigentlich nur der Rückgriff auf die von den Forschern selbst erstellten Publikationslisten, da nur diese eine möglichst vollständige Erfassung aller Publikationen garantieren, aus denen jene ausgewählt werden können, die für die jeweilige Forschungsevaluierung als relevant erachtet werden. Wie Testerhebungen gezeigt haben, ist es allerdings nicht einfach, an aktuelle und vollständige Publikationslisten zu gelangen. An vielen Universitäten, wie etwa an der Universität Graz, gibt es mittlerweile Forschungsdatenbanken, in die alle dort beschäftigten Forscher ihre Publikationen eintragen sollen. Aus diesen Datenbanken lassen sich dann – in der Regel auch von Außenstehenden – auf Knopfdruck Publikationslisten erstellen, die zudem meist einfach elektronisch weiterverarbeitet werden können, was wiederum für die Datengenerierung im Rahmen einer Forschungsevaluierung vorteilhaft ist. Die zunehmende Verbreitung von universitären Forschungsdatenbanken hat dazu geführt, dass viele Forscher keine eigenständigen Publikationslisten mehr führen. Allerdings werden Einträge in die genannten Forschungsdatenbanken im-

mer wieder vergessen oder unrichtig bzw. unvollständig durchgeführt. Zudem sind in solchen Datenbanken häufig nur jene Publikationen enthalten, die ein Forscher während der Beschäftigung an dieser Universität verfasst hat. Soll eine Forschungsevaluierung etwa einen Zeitraum von zehn Jahren umfassen, so müsste man von allen betrachteten Forschern wissen, wann diese wo beschäftigt waren, um entsprechend recherchieren zu können.

Wie aus Tabelle 1 hervorgeht, sind die Datenbanken des Web of Science sowie die Datenbank „Scopus“ als Datenquellen für eine Forschungsevaluierung nur bedingt geeignet, da oftmals nur ein Bruchteil der Publikationen eines Forschers darin enthalten ist. Die Suchmaschine „Google Scholar“ hat sich infolge ihrer unbefriedigenden Suchmöglichkeiten und der weitgehend unstrukturierten Darstellung der Ergebnisse als weitgehend unbrauchbar erwiesen. Um die gefundenen Publikationen zuverlässig den betrachteten Forschern zuzuordnen zu können, würde man die originären Publikationslisten benötigen, wobei man dann aber besser gleich auf diese als Datenquelle zurückgreift (zur Brauchbarkeit von Google Scholar: vgl. Tost/Rindermann 2016).

Vergleichskriterium	Datenquelle	Anzahl (2008-2017)				Durchschnitt
		Forscher A	Forscher B	Forscher C	Forscher D	
Publikationen	Publikationslisten	43	21	29	65	39,5
	Web of Science (% von Liste)	6 (14,0%)	1 (4,8%)	4 (13,8%)	16 (24,6%)	6,8 (17,2%)
	Scopus (% von Liste)	18 (41,9%)	4 (19,0%)	8 (27,6%)	18 (27,7%)	12,0 (30,4%)
Aufsätze	Publikationslisten	10	3	22	15	13,3
	Web of Science (% von Liste)	1 (10%)	1 (33,3%)	4 (18,2%)	11 (73,3%)	4,3 (32,3%)
	Scopus (% von Liste)	4 (40%)	1 (33,3%)	8 (36,4%)	12 (80,0%)	6,3 (47,4%)
Aufsätze-englisch	Publikationslisten	3	2	3	11	4,8
	Web of Science (% von Liste)	0 (0%)	0 (0%)	2 (66,7%)	8 (72,7%)	2,5 (52,1%)
	Scopus (% von Liste)	2 (66,7%)	0 (0%)	2 (66,7%)	8 (72,7%)	3,0 (62,5%)

Tabelle 1: Publikationen nach Datenquelle

Geht man davon aus, dass die verwendeten Publikationslisten der vier betrachteten Forscher vollständig sind, so stehen bei einer Recherche in Scopus im Durchschnitt 69,6% und bei einer Recherche im Web of Science sogar 82,8% dieser Publikationen für eine Forschungsevaluierung nicht zur Verfügung. Beschränkt man sich auf Publikationen in wissenschaftlichen Zeitschriften (im Folgenden als Aufsätze bezeichnet), so reduzieren sich diese Prozentsätze immerhin auf 52,6% bzw. 67,7%. Berücksichtigt man nur englischsprachige Aufsätze, so reduziert sich die „Ausfallsquote“ nochmals auf 37,5% bzw. 47,9%. Dies liegt daran, dass in beiden Datenbanken schwerpunktmäßig Aufsätze in englischsprachigen wissenschaftlichen Zeitschriften erfasst werden (zur Berücksichtigung deutschsprachiger Zeitschriften in diesen Datenbanken: vgl. Clermont/Schmitz 2008).

Generell problematisch gestaltet sich die Datenerhebung erfahrungsgemäß hinsichtlich jener Forscher, die kaum Publikationen aufweisen. Jemand, der in zehn Jahren nur wenige Publikationen vorzuweisen hat, wird eher keine Publikationsliste erstellen. Greift man nun auf die vorhandenen Forschungsdaten-

banken zurück, wird man wenig bis gar nichts finden. Da man in einem derartigen Fall tendenziell vermutet, dass auf die Eintragungen vergessen wurde, wird man versuchen, die betroffenen Personen zu kontaktieren, um die Situation zu klären. Derartige Versuche scheitern jedoch sehr häufig, da kaum jemand offen zugeben wird, dass eigentlich keine Forschungsleistung in Form von Publikationen erbracht wurde.

4 Für eine Forschungsevaluierung relevante Publikationen

Hat man sich nun für eine der oben genannten Datenquellen als Datenbasis für eine Forschungsevaluierung entschieden, stellt sich die Frage, welche der dort gelisteten Publikationen bzw. Publikationsarten für die Evaluierung herangezogen werden sollen. Als mögliche Publikationsarten nennt Stock (2000, 250f.) Buchhandelsmedien, Patente, graue Literatur, audiovisuelle Medien und Internet-Dokumente, wobei die wichtigste Gruppe der Buchhandelsmedien Bücher und Zeitschriften, jeweils in gedruckter als auch elektronischer Form, umfasst. Für den Bereich der Betriebswirtschaftslehre spielen Patente, audiovisuelle Medien sowie Internet-Dokumente – abgesehen von wissenschaftlichen Aufsätzen – praktisch keine Rolle. Graue Literatur sollte nach Ansicht des Autors für eine Forschungsevaluierung prinzipiell nicht herangezogen werden, da es hier in der Regel keine standardisierte Qualitätskontrolle gibt. Bleiben also nur die Buchhandelsmedien.

Die darin enthaltenen Publikationen werden im Zuge der nachfolgenden Ausführungen in folgende Publikationsarten untergliedert:

- a) Aufsätze in wissenschaftlichen Zeitschriften: Das Kriterium der Wissenschaftlichkeit sollte jedenfalls gegeben sein. Indizien dafür könnten etwa vorgesehene Begutachtungsprozesse für eingereichte Beiträge, renommierte Verlage, Institutionen bzw. Personen als Zeitschriftenherausgeber sowie ein gewisser Mindestumfang der veröffentlichten Beiträge sein. Alternativ könnte man sich auf Zeitschriften beschränken, die in einschlägigen Zeitschriftenrankings gelistet sind (z.B. JOURQUAL 3 des Verbands der Hochschullehrer für Betriebswirtschaft) oder die über einen Impact-Faktor verfügen.
- b) Aufsätze in Sammelwerken: Darunter fallen beispielsweise Beiträge in Herausgeberbüchern, Festschriften oder auch Tagungsbänden. Problematisch erscheint bei dieser Publikationsart die oft nur rudimentär vorhandene oder nicht überprüfbare Qualitätskontrolle in Form eines Peer-Review Prozesses.
- c) Monographien (Bücher): Dabei kann es sich um umfassende wissenschaftliche Abhandlungen zu einem Thema und um Lehrbücher handeln. Lehrbücher sollten eigentlich für eine Forschungsevaluierung nicht herangezogen werden, da das Verfassen derartiger Werke eher eine Lehr- als eine Forschungstätigkeit darstellt.
- d) Herausgeberschaft von Sammelwerken: Auch hier stellt sich die Frage, inwieweit es sich bei solchen Konzeptions- und Koordinationstätigkeiten überhaupt um Forschungsleistungen handelt, die in eine Forschungsevaluierung eingehen sollten. Dies zeigt sich auch darin, dass tendenziell als

relevant erachtete Forschungsergebnisse eher weniger in Herausgeberwerken als in einschlägigen wissenschaftlichen Zeitschriften publiziert werden.

- e) Sonstige Publikationen: Darunter werden alle nicht unter a) bis d) fallenden Publikationen in Buchhandelsmedien subsumiert. Dies können etwa Editorials, Rezensionen, Leserbriefe, Lexikoneinträge, Tagungsbesprechungen oder auch Aufsätze in nichtwissenschaftlichen Zeitschriften sein. Aufgrund mangelnder Vergleichbarkeit sollten derartige Publikationen im Rahmen von Forschungsevaluierungen nicht berücksichtigt werden.

Insgesamt ist aufgrund obiger Ausführungen festzuhalten, dass für eine Forschungsevaluierung in erster Linie Aufsätze in wissenschaftlichen Zeitschriften herangezogen werden sollten. So sind auch die Ausführungen in Abschnitt 5 – mit Ausnahme von Tabelle 3, die sich auf sämtliche Publikationsarten bezieht – auf diese Publikationsart beschränkt.

Anhand von Tabelle 2 lässt sich erkennen, wie häufig die eben vorgestellten Publikationsarten in Buchhandelsmedien in den Publikationslisten der vier Forscher der Pilotstudie im Betrachtungszeitraum vorkommen. Demnach spielen Monographien und Herausgeberschaften nur eine geringe Rolle, Aufsätze in Sammelwerken und in wissenschaftlichen Zeitschriften sind dagegen von großer Bedeutung, und sonstige Publikationen werden ebenfalls relativ häufig verfasst. Publikationsarten außerhalb von Buchhandelsmedien scheinen in den Publikationslisten der vier Forscher überhaupt nicht auf. Von der in weiterer Folge vorgenommenen Beschränkung auf Aufsätze in wissenschaftlichen Zeitschriften sind die vier hier betrachteten Forscher in unterschiedlicher Weise betroffen: Forscher C, der vorwiegend in wissenschaftlichen Zeitschriften publiziert hat, profitiert davon, Forscher A und D, die sehr viele Publikationen außerhalb von wissenschaftlichen Zeitschriften vorzuweisen haben, verlieren dadurch leicht, und Forscher B, der fast nur in Sammelwerken publiziert hat, verliert dadurch stark.

Publikationsart	Anzahl (2008-2017)				Durchschnitt
	Forscher A	Forscher B	Forscher C	Forscher D	
Aufsätze in wissenschaftlichen Zeitschriften	10	3	22	15	12,5
Aufsätze in Sammelwerken	25	15	5	23	17,0
Monographien	0	0	1	0	0,3
Herausgeberschaft von Sammelwerken	3	0	1	3	1,8
Sonstige Publikationen	5	3	0	24	8,0
Summe	43	21	29	65	39,5

Tabelle 2: Publikationen nach Publikationsart

5 Gewichtung der für eine Forschungsevaluierung relevanten Publikationen

5.1 Überblick

Während im vorangegangenen Kapitel die Frage diskutiert wurde, welche Publikationsarten für eine Forschungsevaluierung relevant sein sollen, wird nun die mögliche Gewichtung dieser relevanten Publikationen erörtert. Insbesondere wird anhand des Beispiels der vier Forscher gezeigt, inwieweit sich unterschiedliche Gewichtungen auf Forschungsrankings auswirken. Dabei wird stufenweise vorgegangen: Für alle Rankings – Tabelle 3 ausgenommen – werden ausschließlich Aufsätze in wissenschaftlichen Zeitschriften, die in den Publikationslisten angeführt sind, herangezogen. Zudem finden nur Aufsätze mit einer Mindestlänge von 5 Seiten Berücksichtigung, da dies in der vorliegenden Studie als Mindestumfang für einen wissenschaftlichen Aufsatz angesehen wird. Im ersten Unterabschnitt wird gezeigt, welche Auswirkungen die Berücksichtigung von Koautorenschaften auf Forschungsrankings hat. Anschließend wird im zweiten Unterabschnitt zusätzlich der Umfang der relevanten Beiträge und im dritten Unterabschnitt die Bedeutung der Zeitschriften, in welchen die Beiträge publiziert wurden, berücksichtigt.

Davor wird anhand von Tabelle 3, die auf Tabelle 1 basiert, gezeigt, welche Auswirkungen die Heranziehung unterschiedlicher Datenquellen als Datenbasis sowie eine darauf basierende Beschränkung auf Aufsätze bzw. englischsprachige Aufsätze auf ein Forschungsranking haben. Für Tabelle 3 wird jede Publikation – unabhängig von Publikationsart, Koautorenschaft, Umfang sowie der Zeitschrift, in der publiziert wurde – mit einem Wert von 1 gezählt.

Vergleichskriterium	Datenquelle	Ranking			
		Forscher A	Forscher B	Forscher C	Forscher D
Publikationen	Publikationslisten	2.	4.	3.	1.
	Web of Science	2.	4.	3.	1.
	Scopus	1.	4.	3.	1.
Aufsätze	Publikationslisten	3.	4.	1.	2.
	Web of Science	3.	3.	2.	1.
	Scopus	3.	4.	2.	1.
Aufsätze-englisch	Publikationslisten	2.	4.	2.	1.
	Web of Science	3.	3.	2.	1.
	Scopus	2.	4.	2.	1.

Tabelle 3: Forschungsrankings nach Datenquelle

Die Tabelle legt offen, dass die Rankingergebnisse relativ quellenunabhängig sind. Lediglich eine Beschränkung auf Aufsätze führt zu einem Wechsel zwischen erstem und zweitem Rang, da Forscher C eben vorwiegend in Form von Aufsätzen publiziert hat. Da es sich bei diesen Aufsätzen aber größtenteils um deutschsprachige gehandelt hat, führt eine Beschränkung auf englischsprachige Aufsätze wiederum zur ursprünglichen Reihung, bei der Forscher D, der insgesamt am meisten publiziert hat, klar an erster Stelle liegt. Forscher A und B können von den beiden Beschränkungen keinesfalls profitieren, da sie vorwiegend in Sammelbänden publiziert haben.

5.2 Koautorenschaft

Falls Publikationen von mehreren Autoren verfasst wurden, was in den letzten Jahren immer häufiger zu beobachten ist, gibt es mehrere Möglichkeiten, wie diese sogenannten Koautorenschaften im Rahmen einer Forschungsevaluierung berücksichtigt werden können (vgl. dazu auch Stock 2000, 258):

- a) Keine Berücksichtigung: Für jeden beteiligten Autor zählt die Publikation voll, also mit 1. Eine solche Zählung erscheint relativ unfair, da dadurch ein Aufsatz in Summe umso mehr zählt, je mehr Autoren beteiligt sind. Würden sich also beispielsweise 10 Autoren dahingehend absprechen, 10 Aufsätze in Koautorenschaft zu verfassen, würde dadurch jeder Autor 10 Punkte erhalten. Im Gegensatz dazu erhielte jeder Autor im Falle der alleinigen Verfassung eines Aufsatzes, was ungefähr den gleichen Aufwand darstellen würde, sofern jeder Autor bei den 10 Publikationen in Koautorenschaft denselben prozentuellen Beitrag leistet, nur einen Punkt.
- b) Gleiche Berücksichtigung aller Autoren, Punktesumme = 1: Hier wird nach der Formel $1/n$ (n = Anzahl der Autoren) vorgegangen. Dies ist die in der Praxis am häufigsten zu beobachtende Variante, was durchaus berechtigt erscheint, da sie zu einer Gleichbehandlung von Allein- und Koautorenschaft führt und in Publikationslisten bzw. Datenbanken zudem Informationen zum relativen Beitrag einzelner Autoren in der Regel nicht enthalten sind (vgl. auch d)).
- c) Gleiche Berücksichtigung aller Autoren, Punktesumme > 1: Bei dieser Variante wird davon ausgegangen, dass eine Koautorenschaft infolge der notwendigen Koordinierung einen Mehraufwand gegenüber einer Alleinautorenschaft darstellt, wodurch eine Punktesumme größer 1 gerechtfertigt erscheint. Hier wird etwa die Formel $2/(n+1)$ eingesetzt, die dazu führt, dass die Punktesumme mit steigender Anzahl der Autoren zunimmt. Bei 2 Autoren erhält jeder Autor 0,67 Punkte und die Punktesumme beträgt 1,33. Gibt es hingegen 10 Autoren, so erhält jeder Autor 0,18 Punkte und die Punktesumme beträgt 1,81.
- d) Berücksichtigung der tatsächlichen Beiträge der einzelnen Autoren, Punktesumme = 1: In der Praxis ist es eher unwahrscheinlich, dass jeder Koautor tatsächlich den gleichen prozentuellen Beitrag zu einer Publikation leistet. Im Falle ungleicher Teilleistungen wäre es gerecht, die tatsächlichen Anteile zu berücksichtigen. Allerdings kann diese Variante konfliktbeladen sein, da es diesbezüglich durchaus unterschiedliche Wahrnehmungen seitens der Koautoren geben kann. Zudem ist diese Information in einschlägigen Datenquellen kaum enthalten.

Den besten Kompromiss stellt aus Sicht des Autors daher Variante b) dar, die sich auch in der Praxis durchgesetzt hat.

Tabelle 4 zeigt, dass sich die unterschiedlichen Varianten der Berücksichtigung von Koautorenschaften auf die Rankingergebnisse nicht auswirken, wobei Variante d) aufgrund fehlender Informationen bezüglich der tatsächlichen Anteile nicht zum Einsatz kommen konnte, was auch in der Praxis fast immer der Fall sein wird.

Koautorenschaft	Punkte/Rang (2008-2017)			
	Forscher A	Forscher B	Forscher C	Forscher D
Variante a): keine Berücksichtigung	10/3.	3/4.	22/1.	15/2.
Variante b): 1/n	3,8/3.	2,3/4.	15,2/1.	8,6/2.
Variante c): 2/(n+1)	5,5/3.	2,4/4.	17,1/1.	10,1/2.

Tabelle 4: Forschungsrankings auf Basis von Aufsätzen – Berücksichtigung von Koautorenschaften

5.3 Umfang

Eine weitere Gewichtungsmöglichkeit von Publikationen, insbesondere von Aufsätzen, ist jene nach dem Seitenumfang der Beiträge. Dabei ist allerdings zu berücksichtigen, dass es aufgrund von unterschiedlichen Formaten, Schriftgrößen und Layouts zu starken Verzerrungen kommen kann, weshalb die Sinnhaftigkeit dieser Vorgangsweise stark in Frage zu stellen ist. Günstiger wäre eine Gewichtung nach dem Umfang in Wörtern oder noch besser in Zeichen, allerdings sind derartige Daten nur bei elektronischen Publikationen verfügbar. Zu beachten ist auch der Umstand, dass Vergleiche über unterschiedliche Wissenschaftszweige hinweg problematisch sein können. Man denke hierbei beispielsweise an die Schwierigkeiten beim Vergleich von quantitativen, d.h. formalen bzw. formellastigen Beiträgen mit eher textlastigen Publikationen.

Tabelle 5 zeigt die Ergebnisse eines Rankings auf Basis einer solchen Gewichtung nach Seiten, wobei vorhandene Koautorenschaften einmal gar nicht und einmal nach der Formel 1/n berücksichtigt wurden. Beide Varianten führen zu identischen Rankings. Zudem gibt es keine Unterschiede zu den Rankings ohne Berücksichtigung des Umfanges im vorigen Unterabschnitt.

Umfang	Seiten/Rang (2008-2017)			
	Forscher A	Forscher B	Forscher C	Forscher D
Keine Berücksichtigung von Koautorenschaften	93/3.	35/4.	218/1.	180/2.
Berücksichtigung von Koautorenschaften: 1/n	35/3.	28/4.	156/1.	111/2.

Tabelle 5: Forschungsrankings auf Basis von Aufsätzen – Berücksichtigung des Umfanges sowie der Koautorenschaften

5.4 Zeitschriftenrankings

Eine in der Praxis verbreitete Gewichtungsmöglichkeit von Aufsätzen ist jene nach dem „Wert“ der Zeitschrift, in welcher der Beitrag publiziert wurde. Zur Festlegung des „Werts“ einer Zeitschrift können verschiedene Zeitschriftenrankings herangezogen werden. Für den Bereich der Betriebswirtschaftslehre im deutschsprachigen Raum existiert mit dem bereits erwähnten JOURQUAL 3 ein einschlägiges Ranking. Dabei werden die für Betriebswirte besonders relevan-

ten Zeitschriften in die fünf Kategorien „A+“, „A“, „B“, „C“ und „D“ unterteilt, wobei „A+-Zeitschriften“ das höchste Renommee genießen. Offen bleibt dabei, mit welchen Gewichtungsfaktoren die einzelnen Kategorien verknüpft werden sollten. Drei mögliche Varianten, die im Rahmen dieser Pilotstudie eingesetzt werden, sind:

- a) Nur die hochwertigsten Publikationen, nämlich jene in „A+-Zeitschriften“ und „A-Zeitschriften“, werden mit jeweils einem Punkt gezählt.
- b) Alle Publikationen in gelisteten Zeitschriften werden gezählt, wobei für Beiträge in „A+-Zeitschriften“ 10 Punkte, für Beiträge in „A-Zeitschriften“ 6 Punkte und für Beiträge in „B-Zeitschriften“ 3 Punkte vergeben werden; für Beiträge in „C-Zeitschriften“ und „D-Zeitschriften“ wird jeweils 1 Punkt vergeben.
- c) Abwandlung der Variante b) mit einer stärkeren Berücksichtigung höherwertigerer Publikationen, wobei die Punktwerte hier bei 16 (A+), 8 (A), 4 (B), 2 (C) und 1 (D) liegen.

Eine andere Möglichkeit (d)), um den Wert einer Zeitschrift zu bestimmen, ist ein Rückgriff auf den entsprechenden Impact-Faktor dieser Zeitschrift auf Basis der Datenbanken des Web of Science (vgl. dazu auch Stock 2000, 258ff. oder Bauer 2003).

Bei allen vier Varianten stellt sich die Frage, wie mit Beiträgen umgegangen werden soll, die in Zeitschriften erschienen sind, die nicht im JOURQUAL 3 gelistet sind bzw. nicht im Web of Science enthalten sind und daher keinen Impact-Faktor aufweisen. Im Rahmen der vorliegenden Studie wird so vorgegangen, dass Beiträge in solchen Zeitschriften für die Forschungsevaluierung nicht berücksichtigt werden.

Anhand von Tabelle 6 lässt sich erkennen, zu welchen Ergebnissen Rankings auf Basis einer solchen Gewichtung nach der Bedeutung der Zeitschriften, in denen publiziert wurde, führen. Vorhandene Koautorenschaften wurden dabei nach der Formel $1/n$ berücksichtigt. Aufgrund der „Randstellung“ der Forschungsbereiche der vier betrachteten Personen im Hinblick auf die Betriebswirtschaftslehre sind nur zwei Zeitschriften, in denen publiziert wurde, im JOURQUAL 3 gelistet. Somit sind die darauf aufbauenden Rankings nicht aussagekräftig. Anders sieht es hinsichtlich des Rankings auf Basis der Impact-Faktoren aus: Hier hebt sich Forscher D ganz klar von den übrigen drei Forschern ab. Dies liegt vor allem daran, dass Forscher D mit Abstand am meisten englischsprachige Beiträge verfasst hat, was bereits in Tabelle 1 ersichtlich war.

Zeitschriftenrankings	Punkte/Rang (2008-2017)			
	Forscher A	Forscher B	Forscher C	Forscher D
Variante a): JOURQUAL 3: nur "A+ und A-Zeitschriften"	0	0	0	0
Variante b): JOURQUAL 3: 10/6/3/1/1 Punkte	0,8/2.	0/3.	1,5/1.	0/3.
Variante c): JOURQUAL 3: 16/8/4/2/1 Punkte	1/2.	0/3.	2/1.	0/3.
Variante d): Impact-Faktor laut Web of Science	0,8/3.	0/4.	1,4/2.	14,5/1.

Tabelle 6: Forschungsrankings auf Basis von Aufsätzen – Berücksichtigung von Zeitschriftenrankings sowie von Koautorenschaften

6 Resümee

Zusammenfassend lässt sich sagen, dass die vorliegende Pilotstudie Folgendes offenlegt: Eine Forschungsevaluierung, die eine Erstellung von Forschungsrankings auf Ebene der einzelnen Forscher zum Ziel hat, sollte jedenfalls auf die Publikationslisten der Forscher als Datenquelle zurückgreifen. Hinsichtlich der Frage „Was ist eine Publikation?“ lässt sich sagen, dass eine solche Forschungsevaluierung am sinnvollsten auf Basis von Aufsätzen in wissenschaftlichen Zeitschriften durchgeführt werden sollte, da andere Publikationsarten schwer vergleichbar erscheinen. Hinsichtlich der Frage „Was zählt eine Publikation?“ hat die vorliegende Pilotstudie gezeigt, dass es keine optimale Variante für die Gewichtung der einzelnen Aufsätze gibt. Koautorenschaften sollten aber jedenfalls berücksichtigt werden. Ein Rückgriff auf Zeitschriftenrankings ist zwar durchaus sinnvoll, sofern einschlägige existieren. Forschung in Randbereichen oder fachbereichsübergreifende Forschung wird dabei in der Regel jedoch nur unzureichend abgedeckt. Für den Bereich der Betriebswirtschaftslehre bieten sich JOURQUAL 3 oder als Alternative die Impact-Faktoren der Journals im Web of Science an. Die durchgeführte Pilotstudie, die lediglich vier Forscher umfasste, hat gezeigt, dass die Rankings – unabhängig von der gewählten Vorgangsweise – relativ stabil sind.

Erstaunlich ist, dass Stock in seinem Beitrag aus dem Jahr 2000 bereits auf beinahe alle der hier diskutierten Punkte bzw. Varianten eingegangen ist, so dass dieser Aufsatz noch immer eine hervorragende Einführung in die Thematik der Forschungsevaluierung darstellt.

Literatur

- Albers, S. (2015): What drives publication productivity in German business faculties?, in: sbr 67 (Jan.), 6-33.
- Bauer, B. (2003): Habilitationskriterium Impact-Factor. Wie evaluieren medizinische Fakultäten wissenschaftliche Leistungen von Habilitanden?, in: medizin-bibliothek-information 3 (2), 40-43.
- Bogetoft, P./Otto, L. (2011): Benchmarking with DEA, SFA and R, New York: Springer.
- Clermont, M./Schmitz, C. (2008): Erfassung betriebswirtschaftlich relevanter Zeitschriften in den ISI-Datenbanken sowie der Scopus-Datenbank, in: Zeitschrift für Betriebswirtschaft 78 (10), 987-1010.
- Dilger, A./Müller, H. (2012): Ein Forschungsleistungsranking auf der Grundlage von Google Scholar, in: Zeitschrift für Betriebswirtschaft 82 (10), 1089-1105.
- Dorsch, I./Askeridis, J./Stock, W. (2018): Truebounded, overbounded, or underbounded? Scientists' personal publication lists versus lists generated through bibliographic information services, in: Publications 6 (1), 1-9.
- Dyckhoff, H./Thieme, A./Schmitz, C. (2005): Die Wahrnehmung deutschsprachiger Hochschullehrer für Betriebswirtschaft in der internationalen Forschung. Eine Pilotstudie zu Zitationsverhalten und möglichen Einflussfaktoren, in: Die Betriebswirtschaft 65 (4), 350-372.
- Fülbi, R./Weller, M. (2011): A glance at German financial accounting research between 1950 and 2005: A publication and citation analysis, in: sbr 63 (Jan.), 2-33.

- Johnes, G. (1995): Scale and technical efficiency in the production of economic research, in: *Applied Economics Letters* 2 (1), 7-11.
- Meyer, M./Waldkirch, R./Zaggl, M. (2012): Relative performance measurement of researchers: the impact of data source selection, in: *sbr* 64 (Oct.), 308-330.
- Morgan, K. (2004): The research assessment exercise in English universities, 2001, in: *Higher Education* 48 (4), 461-482.
- Reichmann, G. (2007): Evaluierung von Studienplänen. Ermittlung von Möglichkeiten zur Verbesserung von Studienbedingungen, in: *Zeitschrift für Hochschulrecht, Hochschulmanagement und Hochschulpolitik (zfhr)* 6 (2), 37-45.
- Reichmann, G./Sommerguter-Reichmann, M. (2010): Efficiency measures and productivity indexes in the context of university library benchmarking, in: *Applied Economics* 42 (1-3), 311-324.
- Rost, K./Frey, B. (2011): Quantitative and qualitative rankings of scholars, in: *sbr* 63 (Jan.), 63-91.
- Stock, W. (2000): Was ist eine Publikation? Zum Problem der Einheitenbildung in der Wissenschaftsforschung, in: Fuchs-Kittowski, K./Laitko, H./Parthey, H./Umstätter, W. (Hrsg.): *Wissenschaft und Digitale Bibliothek. Wissenschaftsforschung Jahrbuch 1998*, Berlin: Gesellschaft für Wissenschaftsforschung, 239-282.
- Stockmann, R./Meyer, W. (2014): *Evaluation – Eine Einführung*, 2. Aufl., Stuttgart: UTB.
- Tost, C./Rindermann, H. (2016): Evaluation psychologischer und pädagogischer Forschungsleistungen mithilfe der bibliometrischen Datenbanken Scopus und Google Scholar, in: *Zeitschrift für Evaluation* 15 (2), 241-267.
- Wolf, J./Rohn, A./Macharzina, K. (2005): Institution und Forschungsproduktivität. Befunde und Interpretationen aus der deutschsprachigen Betriebswirtschaftslehre, in: *Die Betriebswirtschaft* 65 (1), 62-77.
- Wottawa, H./Thierau, H. (2003): *Lehrbuch Evaluation*, 3. Aufl., Bern: Hans Huber.