Journal of Economic Literature
Vol. XXXVII (March 1999), pp. 157–183

The Young Person's Guide to Writing Economic Theory

WILLIAM THOMSON[1]

1. *Introduction*

HERE ARE MY recommendations for writing economic theory (and, to some extent, giving seminar presentations). My intended audience is young economists working on their dissertations or preparing their first papers for submission to a professional journal.

Although I discuss general issues of presentation, this essay is mainly concerned in its details with formal models. It does not cover the writing up of empirical work. However, since most papers begin with the introduction and the analysis of a model, I hope that it will be useful to anyone, irrespective of field, and not just to fledgling theorists.

The principles of good writing—simplicity, clarity, unity—are universal, but when it comes to putting them into practice, multiple choices are often available, and these recommendations to follow unavoidably reflect my personal tastes. Also, they are occasionally incompatible. This is where judgement comes in. Exercise yours. I make much use of the imperative mode, but I can well imagine that you will come down differently on a number of the issues I raise. What is important is for you to think about them.

Good writing requires revising, revising, and revising again. Undoubtedly, you will spend many months perfecting your first papers, but this work is one of the wisest investments that you will ever make. In your future papers, you will face the same issues again and again, and with the experience you will have gained, you will be able to handle them quickly and efficiently.

Do not think that if your ideas are interesting, people will read your work whether or not it is well written. Your papers are competing with many others that constantly arrive on the desks of the people you hope to reach, so if it is not clear to them fairly quickly that they will get something out of reading your work, they will not even start.

Finally, putting your results on paper is not subsidiary to producing them. The process of writing itself will lead you to

[1] Department of Economics, University of Rochester, Rochester, NY 14627. This is an abbreviated version of a paper entitled "Writing Papers," which is available from the author upon request. I encourage readers to send me their comments at wth2@troi.cc.rochester.edu. I thank Marcus Berliant, Youngsub Chun, Jacques Crémer, John Duggan, James Foster, Tarık Kara, Jerry Kelly, Bettina Klaus, Kin Chung Lo, Leslie Marx, Lionel McKenzie, Philip Reny, Suzanne Scotchmer, and Jean-Max Thomson for their helpful comments and James Schummer for making a reality of my fantasy of letters tumbling down a cliff (footnote 10) in accordance with the laws of physics. I also thank Toru Hokari for the figures. My greatest debt is to John McMillan, Martin Osborne, John Pencavel, James Schummer, the editors, and two anonymous referees of this journal for their numerous and extremely useful suggestions.

new knowledge. Learn to write but also write to learn.[2]

2. *General Principles*

Convey your message efficiently.[3] By leafing through your article, a reader should be able to easily spot the main results, figure out most of the notation, and locate the crucial definitions needed to understand the statement of each theorem.

Readers who have found your central points interesting and want to know more, but have little time to invest in your work, should then be able to get an idea of your methods of proof by visual inspection. It is often quite informative just to glance at the way an argument is structured and to identify the central assumptions and the known theorems on which it is based. Think about the way *you* read a paper. You probably do not proceed in a linear way. Instead, you scan it for the formal results and look around them for an explanation of the notation and terminology that you do not recognize or guess. You do not like having to spend too much time to find what you need. Your readers probably feel the same way about your work.

The Components of a Paper. Your title should be as descriptive of your topic as possible. Devote time to your abstract, as it is on that basis that many potential readers will decide whether to continue. In your acknowledgment footnote, be generous. Include the seminar participant who suggested a name for a condition you introduced, or directed you to a pertinent reference. Your apportionment of credit among the various people who helped you, however, should be commensurate with the time and effort they spent and the usefulness of the suggestions they made. The referee who sent you five pages of comments deserves recognition in a separate sentence.

In your introduction, briefly place your work in the context of the existing literature and describe your main findings. Do not start with a two- or three-page survey of the field; your reader will want to know what your contribution is sooner than that. Use plain language, and skip the technical details. Your literature review should not be a mere enumeration of previous articles. In describing the work on which you build, give priority to the development of the ideas rather than to telling us who did what, although this information should be included, and where you stepped in should be unambiguous. You need not repeat in the body of the paper all of the points that you made in the introduction, although some repetition is unavoidable. On the other hand, I do not generally favor relegating proofs to appendices (more on this later).

Your conclusion should not be a rehashing of the introduction. However, a compact summary of your results and a statement of the main lesson to be drawn from your analysis is a good lead to a list of specific open questions and a general discussion of promising directions for future work, all of which do belong there. In your bibliography, give the relevant background papers. If a good survey is available, mention it. You may have to include papers that you did not use, and papers that you discovered only after you completed yours. Check references carefully, and update them as papers get published.

[2] I owe this formula to William Zinsser's 1989 pedagogical essay. *Writing to Learn,* a book that I strongly recommend.

[3] This paper is longer than the average, but except in Lake Wobegon, not all papers can be shorter than the average. Actually, I do not have a recommendation on how long a paper should be, except for "Make it as long as it needs to be, no longer, and no shorter." If its structure is clear, length by itself is not a problem.

The structure of your paper should be clear, as should the structure of each section, each subsection, and each paragraph. To better see how your paragraphs fit together, summarize each of them in one sentence. Does the string of these sentences make sense? It should. Perform this exercise also at the level of subsections, and then sections.

Show that what you did is interesting and has not been done before. To show that your results are significant, the temptation is great to present them with the utmost generality, with big words, and in gory detail. Resist it! Try instead to make your argument appear simple, and even trivial. This exercise in humility will be good for your soul. It will also give referees a warm feeling about you. Most importantly, it will help you prove your results at the next level of generality.

Because the refereeing process and publication constraints often have the unfortunate effect of wiping out from a paper most of what could make it easily understandable, you may think that if yours does not contain at least one result that looks difficult, it is not ready for submission. You are rightly proud of the sophisticated reasoning that led you to your findings. Nevertheless, work hard to make them look simple.[4]

To show that what you do has not been done before, explain how your assumptions differ from the assumptions used in related literature, and why these differences are significant, both conceptually and technically. Demonstrate your knowledge of this literature by citing the relevant articles and telling us how they pertain to your subject. Also, motivate your work, but do not over-motivate it, or your readers will get suspicious.

Do not forget the process by which you made your discovery. By the time your paper is finished, it will cover an arbitrary number of goods and agents, general production possibilities, uncertainty, and so forth, and nobody will understand it. If you read it several months later, you will not understand it either. You got to your main theorem in small steps, by first working it out for two agents, two goods, linear technologies, and with no uncertainty, and by drawing lots of diagrams. It is also by looking at simple versions of your model that your reader will understand the central ideas, and it is most likely these central ideas, not the details of proofs, that will help her in her own work.

Reproducing the process of discovery in a paper is not easy, but try. In a seminar, quite a bit more can be done because of the informality of the occasion. Explaining how you came to the formulation you eventually chose and to your results, however, is not a license to a rambling discussion in which notation, definitions, assumptions, and motivation are all mixed up, like the ingredients of a big salad. Even worse is adding some semi-formal algebraic manipulations (tossing the salad?), and suddenly confronting us with the sentence: "We have therefore proved the following theorem: . . ." As a reader, I feel as if I have been mugged when I find myself in that situation.

Another principle that has wide validity is that good exposition means going back and forth between the general and the particular, and I will give several illustrations of it.

Learn from your errors. There is nothing like having misunderstood something to really understand it, and

[4] As a young economist, it is natural that you should be proud of the complicated things you achieve; as you get older, you will become proud of the simple things you do. (Of course, it is not because you will not be able to handle the complicated things anymore.)

there is nothing like having seriously misunderstood it to really, really understand it. Instead of being embarrassed by your errors, you should cherish them. I will even say that you cannot claim to have understood something until you have also very completely understood the various ways in which it can be misunderstood. It has been said before, and better: "Erreur, tu n'es pas un mal." (Gaston Bachelard 1938)

Your readers are likely to be victims of the same misunderstandings that you were. By remembering where you had trouble, you will anticipate where you may lose them, and you will give better explanations. In a seminar, quickly identifying the reason why someone in the audience is confused about some aspect of your paper may save you from a 10–minute exchange that otherwise would force you to rush through the second half of your presentation.

3. *Notation*

Choose notation that is easily recognizable. If you have no problem remembering what all of your variables designate, congratulations! But you have been working on your paper for several months now. Unfortunately, what you call x is what your reader has been calling m since graduate school.

The best notation is notation that can be guessed. When you see a man walking down the street with a baguette under his arm and a beret on his head, you do not have to be told he is a Frenchman. You know he is. You can immediately and legitimately invest him with all the attributes of Frenchness, and this greatly facilitates the way you think and talk about him. You can guess his children's names—Renée or Edmond—and chuckle at his supposed admiration for Jerry Lewis.

Similarly, if Z designates a set, call its members z and z', perhaps x, y, and z, but certainly not b, or ℓ. Upon encountering z and z', your reader will immediately know what space they belong to, how many components they have, and that these components are called z_i and z'_j. If Φ is a family of functions, reserve the notation φ and $\tilde{\varphi}$, (perhaps ψ or even f) for members of the family, but certainly not α or m.

If R_i is agent i's preference relation, you may have to designate his most preferred bundle in some choice set by $b_i(R_i)$, his demand correspondence by $d_i(R_i)$, and so on, but dropping this functional dependence may not create ambiguities. For instance, you may write b_i and d_i, provided that you designate agent j's most preferred element in the choice set and his demand correspondence by b_j and d_j, and the comparable concepts when agent i's preferences are changed to R'_i by b'_i and d'_i.

Designate time by t, land by ℓ, alternatives by a, mnemonic notation by mn and so on (and make sure that no two concepts in your paper start with the same letter).

Some letters of the alphabet are used in certain ways so generally in your field that their common interpretation may get in the way of other uses that you want to make of them. You will probably be better off accepting tradition. Do not designate just any quantity by $\in$. Reserve this letter for small quantities or quantities that you will make go to zero.[5] Call your generic individual i, his preference relation R_i, his utility function u_i, and his endowment vector ω_i. The production set is Y. Prices are p,

[5] I like the fragile look of my $\in$, especially when my printer is running out of toner. How could one doubt that the quantity it designates is about to fade into nothingness? However, as a referee reminded me, in econometrics, the error term $\in$ is not necessarily a small quantity, but rather a quantity that one would like to be small.

quantities q. Calligraphic letters often refer to families of sets; so, a is a member of the set A, which is chosen from the family $\mathcal{A}$.

Choose mnemonic abbreviations for assumptions and properties. Do not refer to your assumptions and properties by numbers, letters, or letter-number combinations. Since you state your first theorem on page 10, it will be virtually impossible for us to remember then what "Assumptions A1–A3 and B1–B4" are, but the fact that "Assumptions *Diff*, *Mon*, and *Cont*" refer to differentiability, monotonicity, and continuity will be obvious to a reader starting there. Choose these abbreviations carefully: If you write *Con*, we may not know whether you mean continuity or convexity, so write *Cont* or *Conv*. The cost to you is one extra strike on your keyboard, but your small effort will save us from searching through the paper to find which property you meant. Admittedly, naming each assumption in a way that suggests its content is not always possible, especially in technical fields.

It is common to introduce in parentheses an abbreviation for a condition, next to the full name of the condition at the time it is formally stated. When the abbreviation is used later on, the parentheses are no longer needed.[6]

In axiomatic analyses, many authors refer to axioms by numbers or abbreviations, but I do not see any advantage to that. The argument that numbers and abbreviations save space is not very convincing given that they will not shorten a 20–page paper by more than five lines, and they certainly will not save time for your reader. If you use different typeface for your axioms, which I strongly recommend (for instance italics, or slanted type), each axiom stands out from the text and is perceived globally, as a unit: it is not read syllable by syllable. An alternative way to achieve this important visual separation of the axioms from the text is to capitalize them.

Never use abbreviations in a section heading.

Do not bother introducing a piece of notation if you use it only once or twice. There is no point in defining a new piece of notation if you hardly ever use it. How many times should a concept be used to deserve its own symbol? Three times? Four times? I will let you decide. Certainly, do not bother introducing notation that you never use.

I feel the same way about utility notation when only preferences are involved. It is wonderful, of course, that preference relations satisfying certain properties can be represented by numerical functions, and these representations are sometimes useful or even necessary. But it has become a common excuse to use them even in situations where in fact they only clutter the text. Suppose, for instance, that you want to write that the allocation rule S is ***strategy-proof***. This means that for every agent i, announcing his true preference relation R_i is preferable to announcing any false preference relation R'_i independently of the announcements made by the other agents. Then (here I will skip the quantifications) you can write "$u_i(S(u)) \geq u_i(S(u_{-i}, u'_i))$," but is such an expression preferable to "$S(R)\, R_i\, S(R_{-i}, R'_i)$?" If your paper involves long strings of terms of that form, as may well be the case, utility notation will contribute to an unnecessarily messy look.

Matters are worse if you discuss certain normative issues of welfare economics, social choice, or public finance, because in these fields utility functions have cardinal significance. Even though your theory may only involve the underlying preference relations, some of your

[6] When you begin a proof, write "Proof:" and not "(Proof:)."

readers will come from a different tradition and be tempted to compare utilities, or equate them, or maximize their sum, and so on. On the other hand, if you address some problem of demand theory and you need to calculate matrices of partial derivatives, then of course you cannot avoid utility notation.

Do not define in footnotes important notation that is unlikely to be familiar to your reader, and that you will use in the body of the paper. More generally, do not refer in the main text to terms, ideas, or derivations introduced in a footnote or in a remark, since the reader may have skipped it. There is a hierarchy here that you have to respect.

Save on mathematical symbols. Do not use symbols that are not necessary. For instance, try to avoid multiple subscripts and superscripts. If you have only two agents, call their consumption bundles x and y, with generic coordinates x_k and y_k (instead of x_1 and x_2, with coordinates x_{1k} and x_{2k}). In a text, combinations of subscripts and superscripts look a little better than only subscripts, but in a blackboard presentation, watch out for the sliding superscripts that end up as subscripts. If F is your generic notation for a solution to the bargaining problem, you can certainly refer to the Nash solution as F^N, and when you apply it to the problem (S,d) with feasible set S and disagreement point d, you will get $F^N(S,d)$. But why not simply designate the Nash solution by N? If you can choose the disagreement point to be the origin, as is almost always the case without loss of generality, ignore it in the notation. Altogether, you will calculate $N(S)$, a much lighter expression than $F^N(S,d)$. If you systematically search for such notational simplications, your text will be much cleaner.

Bounds of summation or integration are often (I agree, not always) unambiguous. There is then no need to indicate them. Do not write $\Sigma_{i=1}^{n} x_i$, $\Sigma_{i \in N} x_i$, $\Sigma_i x_i$, $\Sigma_N x_i$, or $\Sigma_{i=1,\ldots,n} x_i$ when, in most cases, Σx_i is perfectly clear. I assure you, upon encountering Σx_i, your readers will be unanimous in assuming that you are summing over i when i runs over its natural domain. Similarly, and although the set consisting of agent i alone should be denoted by $\{i\}$, if you need to refer to it on multiple occasions, you are better off dropping the curly brackets. Do apologize for the abuse of notation though. Similarly, if O designates a list of objects indexed by agents in the set N, you should refer to the shorter list from which the i-th component has been deleted as $O_{N\setminus\{i\}}$, but it has become standard to write O_{-i}. I welcome the shortcut, and I used it earlier. Expressions can be considerably lightened by using such tricks. Imagine that you are on a diet and that each symbol is worth one calorie. You will quickly discover that you can do with half as many. You will improve the readibility of your text and lose weight.

Do not let the reader guess or infer from the context what your inequality symbols mean. Define them the first time you use them. Doing that in a footnote is acceptable.[7] Alternatively, you can give them in a preliminary section of notation.

4. *Definitions*

Be unambiguous when you define a new term. Make it immediately clear that indeed it is new. Do not let your reader think that you may have already

[7] $x \geqq y$ means $x_i \geq y_i$ for all i; $x \geq y$ means $x \geqq y$ and $x \neq y$; $x > y$ means $x_i > y_i$ for all i. You could also use $x \geqq y$, $x > y$, and $x >> y$. It is a very common convention to define these symbols in a footnote, and this is where most of us will look for them when we need them. It is therefore a good idea for you also to define yours in a footnote. Some people have an aversion to footnotes, but personally, I love them. In academic writing, they are often the only place where you will find evidence of life.

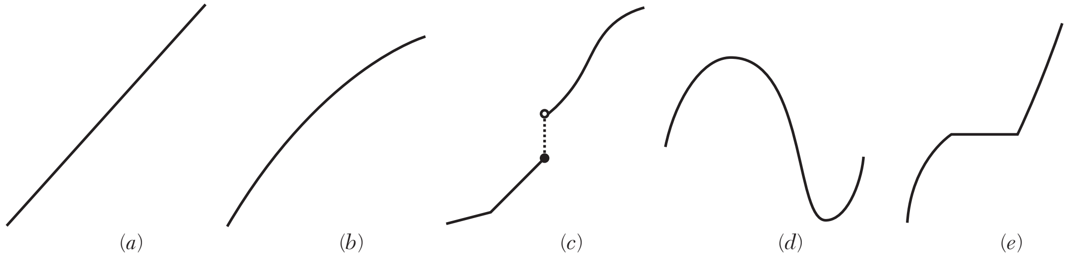


Figure 1. Examples of increasing functions and of functions that are not increasing. (a), (b), and (c): These functions are increasing. (d) and (e): These functions are not.

given the definition but she missed it, or that you are assuming she knows the definition.

Here are three possible ways of introducing a definition: 1. "A function is ***monotone*** if . . ."; 2. "A function is 'monotone' if . . ."; 3. "A function is said to be ***monotone*** if . . ." I prefer the first format and use it throughout this essay, because its phrasing is direct and its different typeface will facilitate its retrieval, if needed. Concerning the typeface, I recommend boldface or boldface italics over italics or plain text between quotation marks, neither of which makes the new terms stand out sufficiently. You should probably display the crucial definitions separately, and you may precede each of them by the word **Definition** in boldface (see the examples below). But do not introduce all definitions in this way, especially if you have many of them, as it will get tedious. Focus on the critical ones.

To avoid repeating quantifications that are common to several definitions, you can group these definitions and state the quantification once: "An allocation rule is ***efficient*** if for all preference profiles R, and all allocations z that it selects for R, there is no other allocation z' that all agents find at least as desirable as z and at least one agent prefers; it is ***weakly efficient*** if instead there is no other allocation z' that all agents prefer to z."

To emphasize certain aspects of your paper, such as important conclusions, exploit the typographical choices at your disposal. Italics is a good one. However, if everything is emphasized, nothing is.

When introducing a novel definition, give illustrative examples. If the definition is a property that an object may or may not have, exhibit: 1. Objects that satisfy the definition; 2. Objects that do not satisfy the definition; 3. Objects that satisfy the definition but almost do not; 4. Objects that do not satisfy the definition but almost do. Examples in Categories 3 and 4 are particularly important as they are responsible for most of the work in the proofs. Conversely, they may be the ones that allow the proofs to go through! In a paper, giving a range of examples that are representative of all four categories is, once again, not easily achieved because of space limitations, but in seminars this can sometimes be done. Here are two illustrations:

Definition. A function f: $[0, 1] \to \mathbb{R}$ is ***increasing*** if for all $t, t' \in [0, 1]$ with $t > t'$, we have $f(t) > f(t')$.

Figures 1a and 1b are dangerous, because they may plant in your reader's mind the seed that you will work with

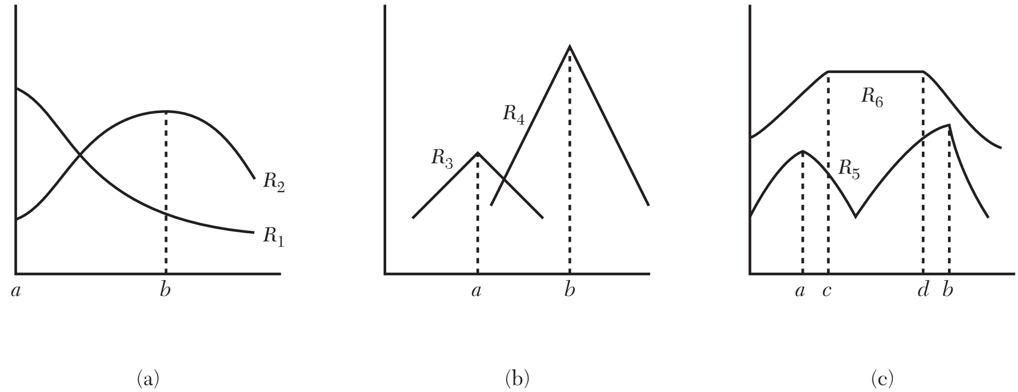


Figure 2. Examples of single-peaked and of non-single-peaked preference relations. (a) These relations are single-peaked, with peaks at a for R_1 and at b for R_2. (b) These relations are single-peaked too, but they are not sufficiently representative of the whole class due to their symmetry. Readers who have not worked with such preferences often assume that symmetry is part of the definition, so you should emphasize that most single-peaked preferences do not have that property. (c) These relations are not single-peaked, since R_5 has two local maxima, at a and b, and R_6 is maximized at any point of the non-degenerate interval $[c, d]$.

functions that are linear, or perhaps concave. Figure 1c is what you need: it represents an increasing function in its full generality, with a kink, a convex part, a concave part, and a discontinuity. Figure 1d is useful too, as it shows a typical violation of the property. Figure 1e is very important because it makes it clear that you want more than that the function be "nondecreasing."[8]

Definition. The continuous preference relation R defined on [0,1], with asymmetric part P, is ***single-peaked*** if there exists $x^* \in [0,1]$ such that for all x, $x' \in [0,1]$ with either $x < x' \leq x^*$ or $x^* \leq x' < x$, we have $x' \, P \, x$.

Figure 2 presents the graphs of the numerical representations of six preference relations. Obviously, R_2 is single-peaked and R_5 is not. But your viewer may not immediately think of R_1 as being single-peaked because its representation achieves its maximum at a corner, or may think that R_6 is admissible, although its representation has a "plateau" and not a peak. You should also make her aware of the fact that you include preferences that do not exhibit the symmetry illustrated in Figure 2b. All of these examples will be very useful to ensure that she fully perceives the boundary of your domain.

Write definitions in logical sequences. Introduce terms in such a way that the definition of each new one only involves terms that are already defined, instead of asking your readers to wait until the end of the sentence or paragraph for everything to be clarified.

For instance, state the dimensionality of the commodity space before you introduce consumers or technologies. In the standard model, a consumer is no more than a preference relation defined over a subset of that space, together with an endowment vector in the space; a technology is simply a subset of the space. In each case, it is therefore natural to specify the space, that is, the

[8] Several readers of this essay objected to sentences such as "this function is nondecreasing," which sounds too much like "this function is not a decreasing function," but means something else. Perhaps we should speak of a "nowhere-decreasing function."

number of goods, first. Therefore, do not write: "$\mathcal{R}_{mon}$ is the class of increasing preferences R, where by increasing is meant that for all $x, y \in \mathbb{R}^{\ell}_{+}$ with $x \geq y$, we have $x\,R\,y$, ℓ being the dimensionality of the commodity space." Instead write: "Let $\ell \in \mathbb{N}$ be the number of goods. The preference relation R defined on $\mathbb{R}^{\ell}_{+}$ is ***increasing*** if for all $x, y \in \mathbb{R}^{\ell}_{+}$ with $x \geq y$, we have $x\,R\,y$. Let P_{mon} be the class of increasing preferences."

As another example, in which $\mathcal{R}^n$ denotes a domain of preference profiles in an n-person economy, do not write:

Definition. The social choice correspondence $F{:}\mathcal{R}^n \to A$ *is* ***Maskin-monotonic*** if for all $R, R' \in \mathcal{R}^n$ and all $a \in F(R)$, if for all $i \in N$, $L(a,R_i) \subseteq L\,(a,R'_i)$, then $a \in F(R')$, where $L(a,R_i)$ is the lower contour set of the preference relation R_i at a, with R and R' being profiles of preference relations defined over A, some alternative space, and Maskin being an economist at Harvard.

Instead write:

Definition. Let Maskin be an economist at Harvard. Let A be a set of alternatives. Given R_i, a preference relation defined over A, and a, an alternative in A, let $L(a,R_i)$ be the lower contour set of R_i at a. The social choice correspondence F: $\mathcal{R}^n \longrightarrow A$ is ***Maskin-monotonic*** if for all $R, R' \in \mathcal{R}^n$ and all $a \in F(R)$, if for all $i \in N$, $L(a, R_i) \subseteq L(a, R'_i)$, then $a \in F(R')$.

Even better, first introduce the basic notation—you will probably use it in other definitions and in the proofs—and only then give the definition. This separation will help highlight the essential idea of the definition.[9] Begin with:

"Let A be a set of alternatives. Given R_i, a preference relation defined over A, and a, an alternative in A, let $L(a,R_i)$ be the lower contour set of R_i at a. Let $\mathcal{R}$ be a class of admissible preference relations defined over A. A ***social choice correspondence*** associates with every profile of preference relations in $\mathcal{R}^n$ a nonempty subset of A."

Now, you can state the definition:

Definition. The social choice correspondence $F{:}\mathcal{R}^n \to A$ is ***Maskin-monotonic*** if for all $R, R' \in \mathcal{R}^n$ and all $a \in F(R)$, if for all $i \in N$, $L(a, R_i) \subseteq L(a, R'_i)$, then $a \in F(R')$.

You may also want to display the hypothesis and the conclusion:

Definition. The social choice correspondence $F{:}\mathcal{R}^n \to A$ is ***Maskin-monotonic*** if for all $R, R' \in \mathcal{R}^n$ and all $a \in F(R)$, if

$$\text{for all } i \in N,\ L\,(a, R_i) \subseteq L(a, R'_i),$$

then

$$a \in F(R').$$

If the hypotheses and the conclusions are simple enough, however, as they are in this example, displaying them may not be needed.

Some will object to the double "if" in the condition as I wrote it, and it does sound awkward. What about replacing the first one with something like "whenever"? Another option is to write: "$L(a, R_i) \subseteq L(a, R'_i)$ for all $i \in N$ implies $a \in F(R')$."

I have seen the recommendation to drop the punctuation at the end of displayed formulas (the hypothesis and the conclusion of the last statement of *Maskin-monotonicity*), but there is far from complete agreement about this. Personally, I prefer all my sentences to be fully punctuated.[10]

[9] Same thing with propositions and theorems: Do not introduce new notation in their statements.

[10] When my daughters were in primary school, I occasionally went to their school to help out with the kids' writing, and my main job was to check that every sentence they wrote began with a capital letter and ended with a period. I have learned this lesson well, and when I see a sentence that does not end with a period, I experience the same queasiness as when I step too close to the edge of an open s

Make sure that the informal descriptions of your definitions match their formal statements. If you write: "A feasible allocation is Pareto efficient if there is no other feasible allocation that all agents find at least as desirable and at least one agent prefers," your formal definition should not be (still using R to denote a preference profile and introducing $\mathcal{P}$ for the set of Pareto efficient allocations): "$z \in \mathcal{P}$ if (i) $z \in Z$ and (ii) for all $z' \in Z$ such that for some $i \in N$, $z'_i \, P_i \, z_i$, there is $j \in N$ such that $z_j \, P_j \, z'_j$." Instead write: "$z \in \mathcal{P}$ if (i) $z \in Z$ and (ii) there is no $z' \in Z$ such that for all $i \in N$, $z'_i \, R_i \, z_i$, and for some $j \in N$, $z'_j \, P_j \, z_j$."

Separate formal definitions from their interpretations. Formal models can often be given several interpretations. It is, therefore, of great value to separate the formal description of your model from the interpretation you intend in your particular application. For example, first write:

Definition. Let $\mathcal{V}^n$ be a domain of n-person coalitional games. A ***solution on*** $\boldsymbol{\mathcal{V}^n}$ is a function that associates with every game $v \in \mathcal{V}^n$ a point $x \in \mathbb{R}^n$ such that $\Sigma x_i \leq v(N)$.[11]

Then explain: "If F is a solution on $\mathcal{V}^n$, v is a game in $\mathcal{V}^n$, and i is a player in N, the number $F_i(v)$ can be interpreted as the 'value to player i of being involved in the game v,' that is, the amount that he would be willing to pay to have the opportunity to play it. Alternatively, it can be thought of as the amount that an impartial arbitrator would recommend the player should receive."

The advantage of this separation is that it will help your reader (and even yourself) discover the relevance of your results to other situations that she (and you) had not thought about initially. To pursue the example I just gave, the theory of coalitional games is also the theory of cost allocation. Some of your readers are interested only in applications, and not in abstract games; others do not care for the applications. You can catch the attention of all by first giving general definitions and then pointing out the various possible interpretations of your model.

Present the basic concepts of your theory in their full generality. You will almost certainly use concepts that are meaningful much beyond the framework of your paper. It is preferable to introduce them without imposing the extra assumptions that you will need to invoke for your analysis. When you explain what a Walrasian equilibrium is, do not assume convexity, monotonicity, or even continuity of preferences.[12] Of course, these properties are relevant when you turn to the issue of existence of these equilibria, but they have nothing to do with the concept of a Walrasian equilibrium itself.

When you introduce a piece of notation, tell your reader what kind of mathematical object it designates, whether it is a point in a vector space, a set, a function, and so on. Do not write "A pair (p,x) is a ***Walrasian equilibrium*** if . . ." Instead, first define the price simplex $\Delta^{\ell-1}$ in the ℓ-dimensional Euclidean space and define the allocation space X. Then, write "A pair $(p, x) \in \Delta^{\ell-1} \times X$ is a ***Walrasian equilibrium*** if . . ." Similarly, do not write "The function φ is ***strategy-proof*** if . . . k," but instead, after having defined the set of possible preference

[11] Here we have a bit of a notational problem as the n exponent to V^n indicates the n-player case, whereas the n exponent to $\mathbb{R}^n$ indicates the n-fold cross-product of $\mathbb{R}^n$ by itself. To avoid it, you could write $V^{(n)}$, but I do not think that the risk of confusion is sufficiently high to justify the parentheses.

[12] Discontinuous preferences are not easy to illustrate graphically, so if you give a graphical illustration of your concept, you probably should present it for continuous preferences.

profiles $\mathcal{R}^n$ (the cross-product of $|N|$ copies of $\mathcal{R}$ indexed by the members of N), and the allocation space X, write "The function $\varphi: \mathcal{R}^n \to X$ ***strategy-proof*** if . . ."

Indicating explicitly the nature of the objects that you introduce is especially important if the reader may not be familiar with them. By writing "A triple $(\pi, x, y) \in \Delta^{(\ell-1)n} \times \mathbb{R}_+^{(m-l)n} \times \mathbb{R}^\ell$ is a ***Lindahl equilibrium*** if . . . ," you help her realize that π has components indexed by agents (these are the Lindahl individualized prices).

By the way, a sequence of elements of X is not a subset of X, but a function from the natural numbers to X. So, you cannot write $\{x^k\}_{k \in \mathbb{N}} \subseteq X$. Nor can you write $\{x^k\} \in X$. Speak of "the sequence $\{x^k\}$ of elements of X," or of "the sequence $\{x^k\}$ where for all $k \in \mathbb{N}, x^k \in X$."

When you define a concept, indicate what the concept depends on. Do not write "The function f is ***differentiable*** at t if blah, blah, blah of t." Since what follows "if" depends on t, you should write "The function f is ***differentiable at t*** (including "at t" in the expression in italics) if blah, blah, blah of t." Then, you can continue and say "The function f is *differentiable* if it is differentiable at t for all t in its domain." A marginal rate of substitution is calculated at a point, so speak of ***agent i's marginal rate of substitution at x_i***. For an example taken from the theory of implementation, speak of a ***monotonic transformation of agent i's preferences at x_i***, and not just of a ***monotonic transformation.***

When you define a new variable as a function of old ones, it should appear on the left-hand side of the equality or identity symbol. If M has already been defined, and M' is introduced next, with a value equal to M, you should write "Let $M' = M$, " and not "Let $M = M'$."

Do not assume that your readers are necessarily familiar with the definitions you use. There is rarely complete agreement on definitions in the literature. Apparently standard terms are often understood differently by different people. Therefore, define the terms you use, even some that you can legitimately assume everyone has already seen. "Core", "public goods," and "incentive compatibility" are examples of terms that are common enough, but define them. The word "rationality" frequently appears in formal developments in game theory without a definition being given. Do not make such a mistake.

Refer to a given concept by only one name or phrase, even if you have several natural choices. Make one and stick to it. Indicate in parentheses next to your definition, or in a footnote, the other terms that appear in the literature. When you first discuss the general idea, you may use different terms in order to vary language and avoid repetitions repetitions, which admittedly do not sound very good, but after you have formally defined the concept and baptized it, only refer to it by its name.

The terms "game", "game form," and "mechanism" are used by different authors to designate the same concept. Choose one. For example write: "A ***game form***[13] is a pair (S,h). . ." You can also write: a "***game form*** (also known as a mechanism)," thereby telling us that your intention is to use the phrase "game form" since it is in boldface italics, but reminding us that the term "mechanism" is also used. You would be confusing us if you wrote "a mechanism (or ***game form***) . . ."

Do not populate your paper with individuals, agents, persons, consumers, and players. One species is enough.

[13] The terms "game" or "mechanism" are sometimes used.

Universal quantifications can be written as "for all," "for any," and "for every"; "given" can also introduce some object taken arbitrarily from some set. I have seen proofs in which all four ways of quantifying were used, and that did not look good. Be careful about "for any." If you write "If for any $x \in X, f(x) > a, \ldots,$" it really is not clear whether you mean "for all x" or "for some x." The terms "preference relation," "utility," and "utility function" are used interchangeably by some authors, but you should not do so. There are important conceptual distinctions here, to which I alluded earlier. Choose language so as to help keep them straight.

In areas where language has not settled yet, you may have several, perhaps many choices. Do not take this as a license to go back and forth between several terms. Instead, seize the opportunity to steer terminology in the direction you favor.

Name your concepts carefully. When you introduce a definition, you need to find a good name for it, a term or a phrase that suggests its content. If you use a multi-word expression, do not worry too much about its length. Your priority is that it should be clear which concept you are designating. In any case, you can also use abbreviated forms of the expressions you chose. A good way of preparing us for an abbreviated expression is as follows: "A feasible allocation is (Pareto)-***efficient*** if there is no other feasible allocation that all agents find at least as desirable and at least one agent prefers." Later on, you can simply talk about "efficient allocations." Unless you use several notions of efficiency, in which case you obviously need to distinguish between them by means of different phrases, the shorter expression is unambiguous and slightly easier to use.

Actually, I do not think that long expressions are much of a problem in a text, as I explained earlier. In a seminar presentation, however, they may be. On these occasions, look for relatively short ones. Alternatively, you can use the long and more descriptive expression a few times, and when you think that the concept has been absorbed by your audience, tell them: "From here on, I will only use the following shorter expression: . . ."

Avoid unnecessary technical jargon. If a function is order-preserving, do not say that it satisfies "order-preservingness"; the name of the property is "order-preservation." I do not like the phrase "one-player coalition," which we use when discussing cooperative games; you may have to speak separately of individual players and of coalitions (sets of two or more players). A theorem is proved by a person, not by a paper: "this result is established by Smith (1978)" is better than "this result is established in Smith (1978)." In common language, "preferring" means what in economese we often call "strictly preferring," and in our dialect we have the phrase "weakly preferring," which does violence to standard English too. In most cases, we can rephrase so as to avoid these conflicts with common usage. When you feel you cannot avoid a conflict, give priority to your statement being unambiguous.

Keeping in mind that a given condition may have different interpretations that depend on the context, choose neutral expressions that cover the various applications over expressions that are too intimately linked to the particular set-up to which your paper mainly pertains. The requirement that an allocation rule be monotonic with respect to an agent's endowment can be seen from the strategic viewpoint; it will make it unprofitable for the agent to destroy some of the resources he controls.

Alternatively, it may be motivated by fairness considerations; the agent should derive some benefit from an increase in the resources he has earned. Instead of phrases taken from game theory or from the theory of fair allocation, however, use a neutral expression such as "monotonicity," (or "endowment monotonicity" if you also discuss monotonicities with respect to other parameters), and let your readers decide which interpretation they prefer.

Designate assumptions by names that help keep the logical relations between them in mind. *Strict monotonicity* should imply *monotonicity*, a condition that in turn should imply *weak monotonicity*. In an axiomatic study, axioms often come in a variety of forms of different strengths. Name them so as to make their hierarchy clear.

Challenge dominant terminology and usage if you find them inadequate. If your paper is a follow-up to someone's published work, as it almost certainly is, do not feel compelled to use the same language if it was not well chosen, even if the writer is a prominent member of the profession. The same comment applies to notation. For instance, why should the adjective "fair" be used to designate allocations that are both equitable *and* efficient, as it was in the early fairness literature? In common language, the term has no efficiency connotation. Refer to "equitable and efficient allocations." The word "endowment" suggests (admittedly, it does not imply) resources that are owned "initially," prior to exchange and production, so the expression "*initial* endowment" is redundant. Just speak of the agents' endowments.[14] The condition of "independence of irrelevant alternatives" that Nash used in his axiomatic derivation of what we now call the Nash solution, is dangerous. I prefer a phrase such as "contraction independence," which is suggestive of the geometric operation that is being performed, without of course allowing us to infer exactly what this operation is, but Nash's expression is no more informative. The reader will decide on her own whether these contractions are irrelevant. "Maskin-monotonicity" is really an invariance condition: it states the invariance of the social choice under certain transformations of preferences—the term "monotonic" is appropriate to describe these transformations—and designating it by a phrase such as "invariance under monotonic transformations" might be a better idea, especially for audiences that are not familiar with the implementation literature. (In general, naming conditions after their authors is not as useful as naming them in a way that suggests their content.) If the length of this alternate expression bothers you, what about "Maskin-invariance"? If you decide to introduce a new phrase, do not forget to also indicate the names that are commonly used.

Of course, the English language was not developed to label concepts of mathematics or economics, but the closer the fit between the concept you have to name and the common meaning of the word you choose, the better. For most of your conditions, you cannot hope to find a short phrase describing without ambiguity hypothesis and conclusion; strike the right balance between compactness and precision.

[14] Besides, if you have to consider changes in the endowment of a player, to find out for instance whether the owner of two left gloves may gain by throwing away one of them prior to entering the market, you will have to make him go from the pleonastic "initial initial endowments" to the oxymoronic "final initial endowments," and whatever benefit he may derive from his clever move will be more than cancelled by the embarrassment of using bad English.

Use technical terms correctly. Do not use the term "vector" unless you will perform vector space operations. If you have in mind a collection of objects taken from some set, the appropriate terms are "lists", "ordered lists," or "profiles." For instance, the notation $(R_1, \ldots, R_n)$ refers to an ordered list of preference relations (or a preference profile), not a vector of preference relations: you will probably not compute $(R_1 + R_2)/2$. On the other hand, it is often appropriate to present a list $(s_1, \ldots, s_n)$ of strategies as a strategy *vector*; for instance, in a game form designed to implement a solution to a public goods problem, a strategy for an agent may be a public good level, and the outcome function may select the *average* of the announced levels. Consumption bundles are usually vectors. You often compute averages of consumption bundles or multiply them by two.

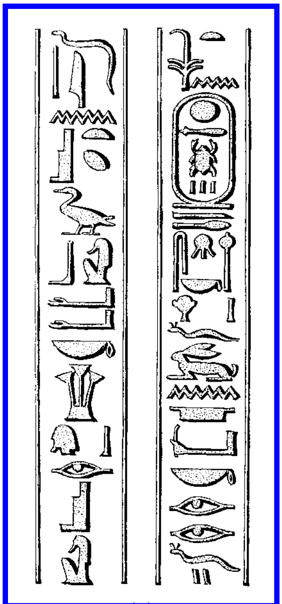
Do not confuse functions with the values they take. If $f:\mathbb{R} \to \mathbb{R}$ is a function, $f(x)$ is the value the function takes when its argument is x. So $f(x)$ cannot be differentiable, or concave, and so on. These are properties of f and not of its values. Designate the function simply by f (this is better than $f(\cdot)$). By the same token, $u_i(x_i)$ is not agent i's utility function; u_i is. Conversely, if u_i is agent i's utility function, it is not also the particular value that this function takes for a certain choice of its argument. If F is a solution to a class of bargaining problems, and S is a problem in its domain of definition, $F(S)$ is not a solution anymore, but something like a "solution outcome," the "solution outcome of S." Alternatively, you can call F a "solution concept" and refer to $F(S)$ as the "solution of S."

Get a good dictionary, and, if English is not your first language, ask for assistance. To weed out from your text gallicisms, nipponisms, sinicisms, and so on, get the help of a native gardener.

5. *Writing Proofs*

Learn LaTeX *or Scientific Word.* One of the first choices you have to make is that of a typesetting software. For your dissertation, I strongly endorse LaTeX, (or TeX, or Scientific Word, whichever one you can handle). LaTeX makes plain text look beautiful, and because it understands the structure of mathematical expressions, its benefits for the writing of mathematics cannot be measured. Moreover, it is widely used (in mathematics, it has truly become the typesetter's LATIN, and you will find it very convenient when collaborating with coauthors dispersed throughout the world. A reader of a previous version of this essay suggested that I recommend the "LaTeX Graphics Companion" of Goosens, Rahtz, and Mittelbach (Addison-Wesley) and "PSTricks" of Timothy van Zandt, advice that was seconded by another reader. If you do not know how to use these softwares, ask one of your younger classmates to teach you (knowledge about computers goes from the young to the old). Also, use a spell-check. When submitting a paper to a journal, respect their style guidelines.

The optimal ratio of mathematics to English in a proof varies from reader to reader, but there is a consensus on a middle range. A proof written entirely in English is often not precise enough and is too long; a proof written entirely in mathematics is impossible to understand, unless you are a digital computer of course. Modern estimation techniques have shown that the optimal ratio of mathematics to English in a proof lies in the interval (52%, 63.5%). Pick the point in that interval that is right for you and stick to it. However, the theorems themselves should be stated



(a)

All I have to do is deduce, from what I know of you, the way your mind works. Are you the kind of man who would put the poison into his own glass, or into the glass of his enemy? . . . Now a great fool . . . would place the wine in front of his own goblet, because he would know that only another great fool would reach first for what he was given. I am clearly not a great fool, so I will clearly not reach for your wine . . . We have now decided the poisoned cup is most likely in front of you. But the poison is powder made from iocane and iocane comes only from Australia and Australia, as everyone knows, is peopled with criminals and criminals are used to having people not trust them, and I don't trust you, which means that I can clearly not choose the wine in front of you . . . But again, you must have suspected I knew the origins of iocane, so you would have known I knew about the criminals and criminal behavior, and therefore I can clearly not choose the wine in front of me.

(b)

Proof: This follows from the inclusion $\varphi \subseteq P$, Part (i) Proposition 1, and Lemma 1 applied to φ. QED

(c)

Figure 3. The ratio of mathematics to English in a proof should be in the interval [52%, 63.5%]. (a) This proof has too much math. Due to the density of mathematical symbols, it is virtually impossible to understand. (I can only make out that it states the existence of ducks having certain properties.) (b) This game-theoretic proof due to William Goldman (1973) has too much English; it is not precise enough and is too long. Not surprisingly, two paragraphs down, the character who produced it is dead. (c) This proof is just right, said Goldilocks, and that is the one she read. It is indeed pleasantly short and clean. Wouldn't you like to know what theorem it proves?

in the simplest English possible. The reader who wants to know more than the probably informal description of results given in your introduction, but does not have much time, will be able to gain a much more precise understanding of your contribution at a very small cost by just reading the theorems. I admit that this is sometimes difficult to achieve, and for technical papers it is probably impossible, but you should try.

Avoid long sentences. A good way to prevent ambiguities is to mainly write one-clause sentences. If English is not your native language, this will also greatly help you avoid grammatical errors. Finally, it will force you to write sentences in logical sequences. Here is an illustration of the idea: "Let (S,h) be a game form. Let $\mathcal{R}^n$ be a class of admissible profiles of preference relations. Given $R \in \mathcal{R}^n$, the triple (S, h, R) is a *game*. A ***Nash equilibrium of (S, h, R)*** is a point $s \in S$ such that for all $i \in N$ and all $s'_i \in S_i$, we have $h_i(s'_i, s_i)\, R_i\, h_i(s)$. If $s \in S$ is an equilibrium, $h(s) \in Z$ is its corresponding ***equilibrium outcome***. Let $E(S, h, R) \subseteq Z$ denote the set of equilibrium outcomes of the game (S, h, R). ***The game form (S, h) implements the correspondence*** $\varphi{:}\mathcal{R}^n \to Z$ ***if*** for all preference profiles $R \in \mathcal{R}^n$, we have $E(S, h, R) = \varphi(R)$.

You may think that your chance for a

Nobel prize in literature will not improve much by this staccato style. Yet I could name several grammatically impaired writers who hardly ever used subordinate or relative clauses and yet got to make the trip to Stockholm! If you really do not like such choppy writing, in your very last draft, reconnect some of your shortest sentences. Similarly, break your text into paragraphs of reasonable size, keeping in mind that too much of a good thing is a bad thing: a sequence of one-sentence paragraphs is not pleasant to read.

A certain amount of redundancy is useful, but do not overdo it. Giving an informal description of the main steps of a proof in addition to the formal proof is not strictly necessary, but it might be quite helpful. Any such explanation, however, should not appear within the proof itself, but outside and preferably before, so as to prepare us for it. The proof itself should be as concise as you can make it without hampering readibility. Similarly, when you state a difficult definition, assist us by giving an informal explanation in addition to the formal statement. Here, too, give it before the formal statement, as this placement will prepare your readers for it. It will also save them[15] frustation: it is indeed annoying to spend time trying to understand a complicated concept when it is first given, only to discover two paragraphs down that the author was willing to help after all.

The same comment applies to figures. If you provided a figure to illustrate a proof, thank you, but why didn't you say so ahead of time, so that we could identify on it the variables as you first introduced them and use it to follow your argument? Warning us of the existence of a figure is especially important because, if your typesetting experience is as limited as mine, you will find it hard to control where the figure ends up (my computer always seems to make those kinds of decisions), and a figure illustrating a particular proof might very well appear on the page that follows the proof instead of next to the proof.

It is often worth explaining very simple things, especially in seminars where you will not have the time to explain the complicated ones in any detail, and especially at the beginning. Indeed, if you lose your audience then, you may have a hard time retrieving it.

After stating an "if and only if theorem," do not refer to the "if part" and the "only if" part, or the "sufficiency part" and the "necessity part." Most people will not know for sure which direction you mean. I have even seen some of the greatest economists being confused about this, and in my personal pantheon, they are people whose approach to economics cannot be described as "literary." Restate the result in each direction as you discuss it. Similarly, would you guess that most of your professors really do not know what a marginal rate of substitution is? But it is true! To most of us, a sentence such as "Agent 1's marginal rate of substitution at z_0 is greater than agent 2's" only means that the two agents' indifference curves through z_0 have different slopes at z_0. We just hope that which is steeper will be clear when we really need to know. Of course, we would never admit it in public, and I most certainly would never put such a confession in writing, for fear of being forever shunned by my colleagues! Instead, compare the

[15] Did you notice that I sometimes refer to "the reader," sometimes to "your reader" (in the singular), sometimes to "your readers" (in the plural), sometimes to "us," your readers? This is an example of an inconsistency of style that should be avoided. Just like this "should be avoided" since I have throughout addressed you, my reader; therefore, I should have written, "that you should avoid." I return to this issue at the end of this essay.

agents' marginal rates of substitution *of good 2 for good 1* at the point z_0; even better, simply talk about their indifference curves being more or less steep at z_0.

It is a great unsolved mystery of neuroscience that someone can prove the fanciest theorems in the most abstract spaces and yet have trouble with some very elementary operations. Remember that. After all, haven't you called your relatives in England when it was 3 *a.m.* there, after having carefully calculated that it would be 3 *p.m.*? You might have failed in such a trivial calculation, and yet brilliantly passed exams where much more of your intellect was being tested.

Use pictures. Even simple pictures can be of tremendous help in making your seminar presentations more vivid. Figures are also very important to lighten a paper, to provide relief from long verbal or algebraic developments, and to illustrate definitions and steps of proofs. Of course, a figure is not a substitute for a proof, and the proof should be understandable without it, but it may give the main idea, and thereby cut by half (probably much more than that, actually) the time your reader will need to understand it. Again, remember the hundreds of little diagrams that you drew on your way to your results.

Label your figures as completely as possible. Label the allocations, the supporting prices, and the endowments. To indicate the efficiency of an allocation, it often helps to shade the upper contour sets in the neigborhood of that allocation. Label a few indifference curves for each agent (some redundancy is useful). If you assume convexity of preference relations and if in fact you draw the indifference curves strictly convex, who owns which indifference curve will be unambiguous. But if you do not make that assumption—you may very well work with linear preference relations or non-convex ones—this ownership will not always be so clear. Avoid unnecessary arrows. You can most often position your labels close to the items they designate without creating ambiguities. Use arrows only if the figure would get too crowded, in particular if the label is too long.[16]

Have one enumeration for each category of objects. Number definitions separately from propositions, theorems, and so on. Some authors use a single list for all of their numbered items, so that for example, Definition 15, which is the tenth definition, is followed by Theorem 16, which is the third theorem, this theorem being followed by Corollary 17, which is the only corollary . . . and so on. Multiple lists are preferable, as they help us understand the structure of the paper. If you have two main sections, with one theorem in each, label the theorems Theorem 1 and Theorem 2. Having a single list certainly facilitates retrieving a needed item, but this benefit is too small. Bringing out the structure of your paper is more important.[17]

State your assumptions in order of decreasing plausibility or generality. When introducing your assumptions, start with the least controversial ones, and write them in order of decreasing plausibility. For utility functions, do not write *A*1: u_i is strictly concave; *A*2: u_i is bounded; *A*3: u_i is continuous. Instead, and here I do not attempt to give names to the conditions, write: *A*1: u_i is

[16] Look at the map of the city where you live—there are hundreds of them—and you will note that all the streets are labeled without arrows and yet without ambiguities! You surely do not need arrows in your figures.

[17] For long documents such as books, adding to the label of a theorem the page number on which it is stated might be useful: Theorem 3.123 is the third theorem of the chapter and appears on page 123.

continuous; $A2$: u_i is bounded; $A3$: u_i is strictly concave.

Introduce your assumptions in related groups. For a general equilibrium model, Assumptions $A1$–$A5$ pertain to consumers and Assumptions $B1$–$B6$ pertain to firms. For a game, Assumptions $A1$–$A3$ pertain to the structure of the game and Assumptions $B1$–$B2$ to the behavior of the players.

Figure out and indicate the logical relations between assumptions and groups of assumptions. If you have many conditions, and many logical relations between them, it is helpful to present these relations in the form of diagrams. The best way to do this is by means of Venn diagrams, each bubble symbolizing the set of objects satisfying one of the conditions.

When you draw two partially overlapping bubbles associated with Conditions A and B, it is because you have identified at least one object satisfying A but not B; at least one object satisfying B but not A; at least one object satisfying both.

You can also use a diagram of arrows and crossed arrows. The advantage of Venn diagrams is that by drawing the bubbles of appropriate size, you can also convey information about the relative strengths of conditions. If A is much stronger than B, draw a much smaller bubble for A. If you prove a theorem under B, whereas A was used in previous literature, your reader will certainly want to know how significant your weakening is. You need to give her some sense of it.

Another advantage of Venn diagrams is that they make it easy to indicate the joint implications of several conditions. If A and B together imply C, the two bubbles symbolizing them intersect within the bubble symbolizing C. With the other technique, you would have to merge two arrows emanating from A and B and point the merged arrow at C. You will end up with a big mess. A disadvantage of Venn diagrams is that for them not to be misleading, you need to figure out *all* of the logical relations between your conditions. But this is another advantage: you need to figure out *all* of the logical relations between your conditions![18] You will not regret doing the work. When you use arrows, by not linking two conditions, you unambiguously indicate not knowing how they are related. That option does not exist with Venn diagrams.

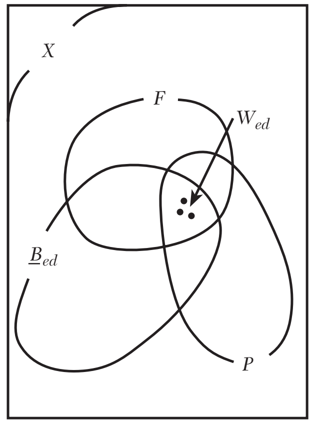
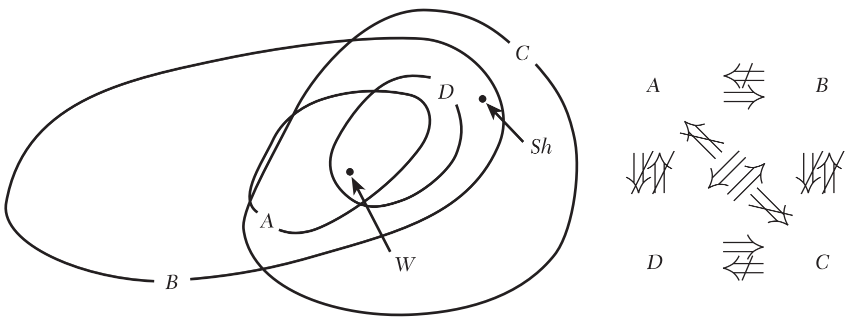


Figure 4. How to indicate logical relations between concepts. Key: X is the feasible set, P the set of Pareto-efficient allocations, F the set of envy-free allocations, $\underline{B}_{ed}$ the set of allocations meeting the equal division lower bound, W_{ed} the set of equal division Walrasian allocations. The set of feasible allocations is so large in relation to the set of Pareto-efficient allocations that its bubble does not even fit in the page. There are continua of Pareto-efficient allocations and of envy-free allocations but typically a finite number of Walrasian allocations. A small tip: breaking the boundary of a bubble to make room for its label is the best way to make unambiguous what is being labelled.

When you use Venn diagrams, you can sometimes draw the bubbles in a way that suggests some of the structure of the sets they designate: if the set is convex, draw a convex bubble; if it is defined by a system of linear

[18] An effective way to do this is as follows: figure out all the illogical relations; what is left are the logical relations.



(a) (b)

Figure 5. Venn diagrams convey much more information than arrows. The two diagrams seem to convey the same information about logical relations, but the Venn diagram (a) allows you to show that "few" objects satisfy condition *A* but not condition *C*, whereas many satisfy condition *B* but not condition *A*. It also allows you to place individual objects, such as the Walrasian rule or the Shapley value, in the appropriate places. (b) I made this diagram of arrows deliberately messy to strengthen my claim that Venn diagrams are more informative than diagrams of arrows, but even if I had been fair, bubbles would have looked better.

inequalities, make its boundary polygonal; if it is a lattice, draw it as a diamond, and so on.

Make sure that there are objects satisfying all the assumptions that you are imposing. Have at least one example. After stating that you will consider economies satisfying Assumptions 1–10, exhibit one that does satisfy all of these assumptions (try Cobb–Douglas; it will probably work). If the class of objects satisfying your assumptions is empty, *any* statement you will make about all of them will be mathematically correct, but of limited usefulness.

Use a common format for the formal statements of your results, and for parts of proofs that are similar. If you have several results that are variants of each other, present them in the same format so as to make their relation to each other immediate. If you first state:

Theorem 1. *If A, B, and C, then D and E*.

do not write your next theorem, which differs from Theorem 1 in that *C* is replaced by *C'* and *E* is replaced by $\tilde{E}$, as

Theorem 2. *Suppose A and B. In addition, consider the class of economies satisfying C′. Then D. Also,* $\tilde{E}$ *holds.*

Instead, use a paralell format:[19]

Theorem 2. *If A, B, and C′, then D and* $\tilde{E}$.

The relation between Theorems 1 and 2 will then be obvious, and your reader will discover it by simply scanning them. By choosing a different format, you would force her to actually *read* their entire statements, and make the comparisons, hypothesis by hypothesis and conclusion by conclusion, that are needed for a good understanding of how the results are related. In some cases, it will be possible to present the two theorems as Parts 1 and 2

[19] This incorrect spelling of paralell (Darn, I did it again!) is an unfortunate consequence of my having finally mastered that of A. Mas-Colell's name (the name for which, in my estimation, the ratio of occurrences of incorrect to correct spellings is the highest in the profession). Do spell names correctly. Dupont does not want to be confused with Dupond any more than Schultze identifies with Schulze. Hernandez and Fernandez are two different people. Thompson is very attached to his "p," and I know for a fact that Thomson has no desire for one.

of a single theorem.[20] Physical proximity and common format are two important ways in which you will facilitate your reader's task.

Similarly, a proof may contain several parts having identical or almost identical structures. Present them so as to make this obvious. Instead of writing Case 1 and Case 2 separately, write Case 1 first, and make sure it is in perfect shape; then copy it and make the minimal adjustments that are necessary to cover Case 2. The similarity of phrasing and format will unambiguously signal to your reader that if she has understood the first part, she can skip the second part. Or if she decides to read Case 2, the marginal cost she will incur will be very small.

Divide proofs into meaningful units, clearly identified. Indent and double indent to indicate structure. Name and number these units: Step 1, Step 2, Case 1, Subcase 1a, Subcase 1b, Case 2, Claim 1, Claim 2. If the proof is sufficiently complex, give each step or claim a title indicating its content. Make sure we know whether this title is a statement that you will prove, or an obvious conclusion that we should reach on our own:

Step 1: The domain of the correspondence φ is compact.

Claim 1a: The domain is bounded. To see this . . .

Claim 1b: The domain is closed. This follows from Lemma 1.

Step 2: The correspondence φ is upper semi-continuous.

If the steps are conceptual units of independent interest, and certainly if they are used in other parts of the paper, as opposed to pertaining to a list of similar cases that have to be checked in turn, call them lemmas (or lemmata, which is the plural form of lemma in Greek; not lemmatas, unless you really have lots of them!), and present them separately. If a proof is long, you may have to number the successive statements that it is composed of. Then, you can refer to them by numbers. Unfortunately, this quickly increases the complexity of the proof, (I mean, how complex it looks). If you do this, only number the essential statements. For instance, if you end a sentence by establishing a statement that is used as a hypothesis in your next sentence, and if the statement is not used elsewhere, you need not number it.

Gather all the conditions needed for a conclusion before the conclusion instead of distributing them on both sides. Hypotheses come first and together. Do not write "If A and B, then D since C." or "If A and B, then D. This is because C." Instead, write "If A, B, and C, then D." Especially for long statements, it helps to visually separate the hypotheses from the conclusions by "then", "we have", or "it follows that". If you write "Since A, B, C, and D," we will not be sure whether you mean "Since A, then B, C, and D," or "Since A and B, then C and D", although technically, the former interpretation has to be the correct one.

Similarly, mathematical statements usually look better when all the quantifications appear together, preferably at the beginning, instead of being distributed on both sides of the predicate. For instance, instead of "For all $x \in X$, we have $x_i > y_i$ for all $i \in N$," write "For all $x \in X$ and all $i \in N$, we have $x_i > y_i$." By the way, this example illustrates a conflict between two of the recommendations that I have made. I just advised to separate mathematical expressions by

[20] Capitalize the word theorem when you refer to a specific theorem, as in Theorem 1 above, but not in a sentence such as "Capitalize the word theorem when . . ." Same rule for propositions, sections, and so on.

English words: "for all $i \in N, x_i > y_i$," does not read as well as "for all $i \in N$, we have $x_i > y_i$," but the formulation "$x_i > y_i$ for all $i \in N$," in which the quantification over agents occurs after the inequality, also achieves the desired separation, and it is shorter.

Be specific about which assumptions, or which parts of assumptions, you need for each step. Do not write "The above assumptions imply that f is increasing" if you need only *some* of the above assumptions to prove that f is increasing. Write "Assumptions 3 and 4 imply that f is increasing". Even better, if you do not need Part (i) of Assumption 4, write "Assumption 3 and Part (ii) of Assumption 4 together imply that f is increasing". Similarly, if Theorem 3 follows from Lemmas 1 and 2, show us exactly how it follows. Do not write "A and B imply C and D," if in fact "A implies C and B implies D." At a very small additional typing cost, you can be much more precise.

When you cite a theorem, be as exact as possible. Refer to a textbook that most of your readers are likely to own or be familiar with. This is especially important for theorems that exist in several forms; we need to know which version you are using. Also, you should probably cite the English edition of a classic text instead of the translated version in your native language, even though that is the one you know well. So write: "By the Brouwer fixed point theorem (Debreu 1959, p.26) . . . Adding the page number is a nice touch.

Verify the independence of your hypotheses. For each hypothesis in each theorem, check whether you could proceed without it. Do not write "Under Assumptions A, B, and C, then D," if A and B together imply C, or if A and B together imply D.

Having put together a toy for one of my daughters, I discovered some parts left in the box. Either these were replacement parts, or I had done something wrong (I will not tell you which, but as a clue, let me say that there never are replacement parts in the box). Similarly, after QED, look in the box for stranded hypotheses. You might have made a mistake, but you might also be pleasantly surprised to find that you can actually prove your theorem without differentiability. Wouldn't you be thrilled if your result applied to Banach lattices (which you did not even know existed two weeks ago), while you thought you were working in boring n-dimensional Euclidean space?

Sometimes, you will be unable to show that a certain hypothesis is necessary for the proof *and* unable to conclude without it either. This is an uncomfortable situation that should keep you up late at night.

A given hypothesis may be the conjunction of several more elementary ones. Then, try to proceed without each of its components in turn. For instance, if you have shown that "Under compactness of the set X, conclusion C holds," do not only check that without compactness, C might not hold anymore. Instead, ask whether "Under boundedness of X, C holds" and whether "Under closedness of X, C holds."

Explore all possible variants of your results. If you prove that "A and B together imply C," do not limit yourself to that statement. Find out whether similar statements hold with A replaced by the closely related conditions A', A^0, and $\tilde{A}$, or B replaced by B' and B^*, or C replaced by C^0. Knowing statement P is not enough. Discover as many statements as possible that are close to P and are also true, and statements that are close to P but are not true. It is as useful to understand the multiplicities of statements around the one you are proving that could be true but are not, as the statement that you are proving. It may even be more useful. Comment on the main variants of your theorem but keep to yourself the least significant ones.

Do not leave (too many) steps to the reader. Give complete arguments. Some steps in a proof may involve standard manipulations and detract from your main point. Perhaps they should not be in the body of the paper, but in an appendix. Do not take them out though. Your reader may not be familiar with a derivation that *you* have seen and performed hundreds of times. Just having the option of assessing the length of a step and recognizing the names of familiar theorems on which it is based will be helpful to her in checking her understanding of the logic of your argument, even if she does not actually read all the details. In general, I do not like too much of the work to be relegated to appendices. When I first look at a paper, I skip most of it anyway, and if I decide to study it more seriously, I find it annoying to have to go back and forth between the body of the paper and the appendix.

If you think a step is obvious, look again. Do not think that your errors necessarily occurred in the hard parts of your proofs (I should say, what you think are the hard parts of your proofs). They may very well have hidden in (what you think are) the easy parts, taking advantage of your overconfidence. After completing your paper, search for the "clearlys" and "obviouslys" and make sure that what you claimed was clear and obvious is, if not clear and obvious, at least true.[21]

Numerical examples are not always useful. It is commonly thought that numerical examples provide easy introductions to complicated proofs. This is true only if the examples are well chosen. A general algebraic expression has in fact the advantage of reminding us of the logic of an argument. If, to fix ideas, you choose $x_1 = 1$ and $x_2 = 8$, the number 9 will refer to the sum $x_1 + x_2$ but it might be helpful to remember this origin: so write "$1 + 8$" instead, or "$9\,(= 1 + 8)$". The expression $x_1 + x_2$ is often preferable. In a three-player game, write the number of coalitions as $2^3 - 1$; we do not care much if that number is equal to 7.

Also, by using numerical examples instead of algebraic notation, you lose track of units of measurement. It makes it harder to check the correctness of expressions.

When you vary a parameter, as a result of which agent 1's income goes from 5 to 7 and agent 2's income from 8 to 5, it will soon be difficult to remember which ones are the initial incomes, which ones are the final incomes, and whose income is 5 and when. If you use well-chosen algebraic notation, for instance by calling the incomes I_1 and I_2 before the change and I'_1 and I'_2 after the change, your reader cannot be confused.

If you insist on using numbers, choose them so that whatever operations you perform on them do not turn them into monsters. If you will divide x_1 by 2, choose x_1 even; if you will take its square root, do not choose $x_1 = 10$. Actually, I take this back. It depends: if the incomes are 5 and 7 initially, and they are cut in half, they will be 5/2 and 7/2 after the change and the fractions will make it easier to remember that they are the new ones. If they were even, you would be tempted to perform the division to get integers and again, the new incomes would be hard to tell apart from the old ones.

In filling a payoff matrix, take all

[21] Do not deduce from this, however, that simply deleting the "clearlys" and "obviouslys" will necessarily eliminate all of your errors.

payoffs to be integers between 0 and 9 so that you do not need to separate them by commas. In each cell of the payoff matrix you can also place the payoffs of the row player slightly higher than that of the column player.

More useful than numerical examples are examples with a small number of agents, a small number of goods, and no production. Then you can save on subscripts, you can use an Edgeworth box, and in your proof you can appeal to the intermediate value theorem instead of to a general fixed point theorem. By the same token, general arguments are sometimes easier to understand than their applications to special situations: it is more transparent why a competitive equilibrium is Pareto efficient when the proof is presented in the general case than for a Cobb–Douglas example, say. There is indeed little to be learned from the calculations for a special case.

Similarly, illustrating a general phenomenon by means of a perhaps incompletely specified geometric example is more informative than a complete argument based on a particular numerical example. The reason is that it may be hard to identify which features of the numerical example are essential to the phenomenon. For instance, to prove that in an Edgeworth box economy there could be several Walrasian equilibria, an example in which preferences are suggested by means of a few indifference curves for each of the two agents suffices. Of course, a few indifference curves do not constitute a preference map, and you have to rely on your readers' experience with such maps for them to mentally complete your figure, or convince themselves that the completion can be done. The alternative is for you to give entire maps, which in most cases will be by means of explicit numerical representations for them. These representations will often be quite complicated, and although they will prove your point beyond doubt, I strongly believe that they will hamper the understanding of the circumstances under which multiple equilibria occur.

If you want to name your agents, do it in a way that helps. When you think numbering your agents from 1 to 4 is too dry in describing an example, try real names, but choose them carefully so as to make it easy to remember who is who. Naming them Bob, Carol, Ted, and Alice will be cute but may be counterproductive. Ted most certainly does not belong in this group. Also, they should be ordered alphabetically: Alice, Bob, Carol, and Dwayne are your four consumers.

In honor of a favorite writer, I have long wanted to call agents 1 and 2 Qfwfq and Xlthlx, but which is actually easier to remember, that agent 1 is endowed with good 1 and agent 2 is endowed with good 2, or that Qfwfq is endowed with apples and Xlthlx endowed with oranges?

By the way, in a seminar, avoid cultural references that are obscure to too large a fraction of your audience, but by all means, do not avoid cultural references altogether because you fear that some of your audience may not understand them. Sometimes it will not be easy to decide. Do you think that in order to prevent those of my readers who don't know French from feeling excluded, I should have resisted the temptation to quote "Erreur, tu n'es pas un mal," thereby depriving the others of this beautiful maxim? Which of the criteria of social choice theory is the right one here?[22]

[22] Once, I referred to Bob and Carol, Ted and Alice in a seminar in which I discussed matching theory, and a member of the audience commented that I was showing my age! I was unfortunately not quick enough—showing my age once again—to

Do not collapse two or three simi_ lar statements into one by indicating the variants in parentheses. Consider the following definition: "The function $f:\mathbb{R} \to \mathbb{R}$ is decreasing (increasing; non-decreasing) if for all x, $y \in \mathbb{R}$ with $x > y$, $f(x) < f(y)$ (respectively $f(x) > f(y)$; $f(x) \geq f(y)$).

The only way for us to be sure we understand this triple definition is to read it three times (once for decreasing, once for increasing, and once for non-decreasing), and yet it is pretty simple. More complicated statements in that format require a mental gymnastics that will unnecessarily exhaust us. Just restate the complete sentence in the various forms you need. I also have a lot of trouble with "and/or" (or is it "or/and"?).

Do not start a sentence with a piece of mathematical notation. Journal editors will red-pencil you if you do, and I agree with them that it does not look good, especially if the notation is lower case. "x designates an allocation" is not pretty. "I is the set of individuals" is not as bad because I is uppercase (but what a grammatical provocation!). "Let x designate an allocation" is what editors will prefer.

Be consistent in your writing style. Do not switch back and forth between first person singular, first person plural, and passive forms. If you write: "In section 3, *I show* that an equilibrium exists. In Section 4, *we establish* uniqueness. To prove these results, *it is assumed* that preference relations are strictly convex and have infinitely differentiable numerical representations. For the proof of the main theorem, *one appeals* to the Brouwer fixed point theorem. *Section 5 concludes.*" your readers will think you need psychiatric help. Are you "I" or "we"? Is it because these assumptions are embarrassing that you suddenly hide behind the passive form? Believe me, we have all made embarrassing assumptions. And why do you let Section 5 conclude when you did all the work? The passive form is found awkward by me and our advice here is to have it replaced. "I" is perhaps too personal. Between "I" and "we", I choose "we", but if you choose "I," we will respect your choice.[23]

Similarly, do not travel back and forth between present and future tenses. Do not write: "First, *I prove* existence. Then *I will apply* the theorem to exchange economies. *I conclude* with open questions." In most cases, using the present tense throughout, even in describing past literature, is just fine.

Choose the sex of your agents once and for all. Flip a coin. If it is a boy, rejoice! If it is a girl, rejoice! And don't subject them to sex change operations from paragraph to paragraph.[24] Two-person games are great for sexual equality. Make one player male and the other female. This will actually facilitate talking about the game and help your reader keep things straight. It will also save you from the awkward "he or she," "him or her,", "his or her"! Alternatively, you may be able to refer to your

reply that by understanding that I was showing my age, and remarking on it, he was showing his! He was right though. I recently asked the students in my graduate class whether they understood the allusion. Not one of them did. And yet, "Bob & Carol & Ted & Alice" (it's a movie) came out only yesterday (30 years ago, to be precise)! From now on, I will use this example only when I give lectures in retirement homes.

[23] As a reader, I rather like the "I" form, which is more engaging, but I am not comfortable using it in formal papers. I use "I" here only because of the informal style that I chose for this paper. Paradoxically, the "we" form is less obstrusive than the "I" form. "We" can also be interpreted as "you and the reader," whom you are taking along, but then be careful if you refer to "our previous work".

[24] For a book, alternating between male and female between chapters might be acceptable though.

agents in the plural, or choose one of them to be a firm, and refer to it as "it."

Be consistent in your choice of running indices. If $N = 1, \ldots, n$, do not write interchangeably "for all $i \in N$," or "for all $i \in 1, \ldots, n$," or "for all $i = 1, \ldots, n$." Pick one formula and stick to it. In most situations, the quantification on the set of agents (to take an example) is clear. Skip it and write "for all i." This helps keep down the density of symbols. In general though, it is good to indicate membership explicitly. For instance, instead of "There exists z for which . . .", write "There exists $z \in Z$ for which . . . ." Therefore, for consistency of style and esthetic reasons, when everything else is explicity quantified, it bothers me a little not to see membership indicated for the set of agents, even if it is pretty obvious that they come from N and not from Mars. So instead of "For all i such that . . . ," I would write "For all $i \in N$ such that . . ."

Do not put quantifiers in the middle of a sentence in English. A sentence such as

"Blah, blah, blah, $\forall x$ such that $P(x)$, blah, blah, blah $\exists y$ such that $Q(x, y)$ and blah, blah, blah."

does not look good. Write "for all" and "there exists." If the mathematical statements introduced by the quantifiers are complex enough, pull them out from the text in English and display them on separate lines, as follows:

"Blah, blah, . . . , blah, blah,

$\forall x$ such that $P(x)$, $\exists y$ such that $Q(x, y)$,

and blah, blah, blah."

The quantifications should always be unambiguous. Remember also that taking the negation of a properly written mathematical statement, with no hidden quantifications, is a trivial operation.

The only mathematical symbols that do not bother me in a text in English are $\leqq$, $\subseteq$, and $\in$, (and the other symbols of the same kind such as the strict inequalities, the strict inclusions, the preference statements, . . .), read as prepositions or verbs.

"Blah, blah, blah, since $x \leq y$, and $x \in A$, and therefore, blah, blah, blah, f is continuous," is fine.

$\exists$ situations where it is convenient to quantify once and $\forall$[25]. For instance, you can open your proof by stating: "In what follows, S denotes an arbitrary element of Σ." Then the requirement that the function $F{:}\Sigma \rightarrow \mathbb{R}^2$ satisfies "for all $S \in \Sigma$, $F(S) > 0$ " can simply be written as:

Positivity: F(S) > 0.[26]

Indicate the end of proofs clearly. Use QED (for *quod erat demonstrandum*), or Halmos' □ (I suppose, for *quod erat quadrandum*.[27]). Delete

[25] See the problem with starting a sentence with a piece of mathematical notation (Section 6.10)! When I wrote earlier that you should not put quantifiers in the middle of a sentence in English, I should have said: do not put them anywhere in such a sentence.

[26] Or "$F > 0$". By the way, do not place your footnote markers at the end of mathematical expressions, as they will look like exponents. Placing them beyond the punctuation mark, as the typographical convention requires, and as I have done here, helps, although logic would sometimes dictate that the marker be attached to a word inside the clause (or the sentence) that ends with the punctuation mark. Compare the marker for this footnote with the marker for the previous one: the position of that earlier marker did not create any ambiguity, as I am sure that you did not think that my intention was to raise the universal quantifier to any power; still it did not look pretty. The same problem arises with quotation marks. I just wrote "$F > 0$". The rule is to write "$F > 0$." This is in agreement with logic if you think of the whole sentence, including the period that ends it, as being the unit that is being discussed. In other contexts, it may be the requirement "$F > 0$" that is under discussion but here, and given that quotation marks look a little like double prime, I admit that placing them after a needed punctuation mark is better, so refer to the requirement "$F > 0$," which is proved in Section 2.

[27] *Circulus?* What about a little circle to indicate the beginning of a proof, matching the little square that closes it?

the redundant "This completes the proof," which precedes □ in your current draft.

6. *Conclusion*

If you follow all of the above recommendations, not only will you be pleased with yourself, your seminar audiences enlightened, your classmates impressed, your parents proud of you, and you will land a job in a top-five department, but most importantly, your adviser will be happy. I readily admit that each of them does not amount to much. However small imperfections, when added together, will take your paper over the line that separates those that can be understood from those that cannot. An Archimedian principle is at work here. You will lose your readers or your seminar audiences much earlier than necessary. In fact, you too will be confused.

Do not fool yourself: very few of your readers will take the time to understand your whole paper, and a large fraction of your seminar audience will not have the faintest idea of what you are talking about when you are half-way through. So, every bit will help in keeping the attention of a few a little longer.

If you are used to certain notational conventions, or terminology, or ways of structuring a proof, they almost necessarily seem the best to you, and perhaps the only ones worth considering. You have to be open-minded and genuinely experiment with other formulations. Only then can you decide what is truly best. The first few times you use a new piece of notation or a new term or a new format, it will appear strange to you. Give it a chance.

Let time elapse beween revisions. If your paper is so familiar to you that you essentially know it by heart, you will never discover your mistakes. You need to let it sit in a drawer for a while. When you pick it up again, it will have a freshness that will allow you to see immediately where it can be improved.

Good writing requires rewriting, and rewriting again. When after many drafts, your paper has become like a smooth and shiny pebble that fits snugly in the palm of your hand, treat yourself to a box of Belgian chocolates. And if you have found these recommendations useful, please save me one!

7. *Related Literature*

As I started circulating this paper, several readers gave me references to similar pedagogical essays written by mathematicians. I am happy to report that their recommendations are not always in contradiction with mine. I found Nicholas Higham (1993) particularly helpful. Paul Halmos' essay in the Norman Steenrod et al. (1983) volume is often cited and deservedly so. Leslie Lamport's (1986) manual is beautiful. (I will even consider forgiving the author for his maxim "All axioms are dull.") William Strunk and Edmund White (1979) is a well-known general manual of style. The *Merriam-Webster Dictionary of English Usage* is an invaluable source, and I am quite fond of the *American Heritage Dictionary of the English Language*. An example of a beautifully written text is the monograph by Gérard Debreu (1959).

References

American Heritage Dictionary. 1992. Houghton Mifflin Company, Third Edition.

Bachelard, Gaston. 1938. *La Formation de l'Esprit Scientifique*.

Chicago Manual of Style. 1993. Fourteenth edition. Chicago and London: U. Chicago Press.

Debreu, Gérard. 1959, *Theory of Value*. New York: Wiley.

Goldman, William. 1973. *The Princess Bride*. Balentine Books.

Higham, Nicholas J. 1993. *Handbook for the Mathematical Sciences*. Philadelphia: Society for Industrial and Applied Mathematics.

Lamport, Leslie. 1986. *LaTeX*. Addison-Wesley Publishing Company.

Merriam-Webster's Dictionary of English Usage. 1994. Springfield, MA: Merriam-Webster Inc.

Steenrod, Norman E.; Paul R. Halmos, Menachem M. Schiffer, and Jean A. Dieudonné. 1983. *How to Write Mathematics*. Rhode Island: American Mathematical Society.

Strunk, William and Edmund B. White. 1979. *The Elements of Style*. New York: Macmillan Co.

Zinsser, William. 1988. *Writing to Learn*. Harper and Row.

This article has been cited by:

1. Ananish Chaudhuri. 2010. Reflections of a journal editor. *New Zealand Economic Papers* **44**:3, 211-215. [CrossRef]