Notational and Terminological Conventions by Peter P. Wakker[1]

Peter P. Wakker
Erasmus School of Economics, Erasmus University, P.O. Box 1738, Rotterdam, 3000 DR, the Netherlands; +31-(0)10 – 408.12.65 (O), +31-(0)10 – 408.12.78/12.64 (S);
Wakker@ese.eur.nl
August, 2024

The primary purpose of this note is not to present general (objective) rules of terminology and notation, and there will be only few of those. This note instead focuses on subjectively chosen conventions of mathematical terminology and notation by the author in situations where several conventions are conceivable. Several conventions will not work well for fields not close to mine (specifed in the beginning of §1). Readers can choose the sections below that may interest them. It is convenient for a field if there are uniform conventions and I hope that this note can contribute to improving such uniformity.

Still a general rule: It is important that terms are short and efficient. If, instead of the term rank-dependent utility, one uses the term "expected utility with rank-dependent probabilities" (as was once proposed), then the intractability of that long term alone amounts to a death sentence for the theory. No-one wants sentences with such clumsy expressions. One of the worst examples of inefficient terminology is the term "multiple choice list." Here the term "multiple" adds zero content to "list", but more than doubles the length. Efficiency score $-\infty$.

1 General notation for preferences

I study individual decision theories, mostly decision under uncertainty. I work in the revealed preference approach central in economics. The generic term for the objects

---

[1] Some colleagues requested that I write this note for a book planned by them.

to be chosen is *prospect*, the term used here, or *option* or (*choice*) *alternative*. Prospects are often lotteries or acts, but can also be income streams, commodity bundles, welfare allocations, medical treatments, objects, and so on. The set of prospects considered is, in general, denoted X, but can be different in particular contexts. It can be capital F (or its script capital $\mathcal{F}$) to denote a set of acts if acts are functions denoted f. I sometimes consider a *choice function* denoted C, that assigns nonempty subsets $C(A)$ to subsets A of prospects.

In most of my papers I do not consider general choice functions, but a *preference relation* $\succcurlyeq$ over prospects, interpreted as binary choice. Preferences are interpreted in the revealed preference sense, as binary choices, no more and no less. Notation $\succ, \preccurlyeq$ $, \prec, \sim$ is as usual. I prefer the symbol $\succcurlyeq$ to $\succsim$. V is *representing* if V maps prospects to the reals with $V(x) \geq V(y)$ if and only if $x \succcurlyeq y$. I then also say that $\succcurlyeq$ *maximizes* V. I avoid calling V a utility function because I use that term for ingredients in representing functionals such as expected utility. I abbreviate expected utility as EU, not EUT for expected utility theory. To keep things short.

Based on long experience, always working with weak orders, I recommend taking weak preference $\succcurlyeq$ as primitive, and not strict preference $\succ$. As I recommend defining preference properties in a weak and not in a strict sense. For example, in my formal definition of risk aversion, expected value maximization and risk neutrality are part of risk aversion. So, I do not define risk aversion in a strict sense. This terminology has some linguistically drawbacks, but the pros outweigh this drawback. Big pro: properties of $\succcurlyeq$ are then often preserved under limit taking.

2 Product sets

The set of prospects is often a product set, mostly with finitely many coordinates, $X_1 \times \cdots \times X_n$. Its elements are then denoted $x = (x_1, \ldots, x_n)$, being n-tuples, called vectors if there is linear-space structure. In some fields a vector x is bold printed, $\boldsymbol{x}$, or the vector has an overbar, $\overline{x}$, but I never did so and felt it was too heavy for my needs. I then always use subscripts only to refer to coordinates and nothing else. Other indexes are denoted as superscripts. Confusions with powers and exponents never happened in my works. So, x is an n-tuple and $x_i \in X_i$ is the ith *coordinate* of x. Both

i and x_i can be called *coordinate*. Further, x^1 and x^2 are two n-tuples, and x_i^j is the ith coordinate of the jth n-tuple x^j.

The sets X_i are *coordinate sets* (or factors). If they are endowed with relations $\succcurlyeq_i$, then:

- *weak monotonicity*: $x \succcurlyeq y$ whenever $x_i \succcurlyeq_i y_i$ for all i;
- *strict monotonicity*: $x \succ y$ whenever $x_i \succ_i y_i$ for all i;
- *strong monotonicity*: $x \succcurlyeq y$ whenever $x_i \succcurlyeq y_i$ for all i and furthermore $x \succ y$ if $x_i \succ y_i$ for one or more i.

If the X_i are subsets of $\mathbb{R}$, then monotonicity definitions often implicitly assume $\succcurlyeq_i = \geq$ for all i. It should be understood that the monotonicity conditions are considerably less plausible if the relations $\succcurlyeq_i$ are not objective, in which case the conditions involve nontrivial weak separability assumptions. I may use the unqualified term monotonicity for one of the above conditions if no confusion will arise.

If $X_1 = \cdots = X_n = C$, i.e., all coordinate sets are identical, then I may sometimes want to denote an element of C without committing to a coordinate. I then use Greek letters, α, β, etc. This notation is useful for instance if I want to let α appear on two different coordinates. By $x_{-i}\alpha$ or, preferably if possible, $\alpha_i x$, I denote the n-tuple x with its ith coordinate replaced by α. By $y_i x$, or $x_{-i}y_i$, I denote the n-tuple x with its ith coordinate replaced by y_i. The difference between $\alpha_i x$ and $y_i x$ should be clear from the context. The notation $\alpha_i\beta_j x$, $x_{-i,j}\alpha\beta$, and $x_{-i,j}y_i z_j$ is similar. For $E \subset \{1, \ldots, n\}$, x_E denotes the restriction of x to E, and $y_E x$ is like y on E and like x elsewhere. If $X_1 = \cdots = X_n = C$, then α is identified with $(\alpha, \ldots, \alpha)$. It gives the notation $\alpha_E\beta$, and the preference relation $\succcurlyeq$ over prospects then gives a preference relation $\succcurlyeq$ over C.

3 Decision under uncertainty and risk

I use the term *risk* for known, objective, probabilities. *Uncertainty* is a catch-all term, referring to cases with completely or very unknown probabilities, somewhat unknown probabilities, but also to cases with known probabilities. Thus, risk is a special case of uncertainty. *Ambiguity* describes the *difference* between uncertainty and risk. Whether a context is risk is exogenously determined, depending on information available, and not endogenously. All applications of the Ellsberg two-urn paradox take the known urn as neutrality point for ambiguity, based on the information we have about the urn and prior to having observed any preference. That is, it is exogenously determined.

Finally, a discussion of the term risk. I do not use the term risk when probabilities are not objective but subjective or a-neutral. Thus, I do not use the concept of source-dependent risk attitude, occurring in some economics papers. While at first the concept works well for nonspecialists, I think it will not work well in the long run. I would not know then how to distinguish ambiguity attitudes and source-dependent risk attitudes on the basis of revealed preference, and work on ambiguity would get complicated. One would then need an extra term for ambiguity neutrality and/or known objective probabilities. I do use the term source-dependent uncertainty attitude. The case of known objective probabilities is important. Why not have a specific short term for it? Agreeing with most common current conventions in economics. One should be warned that other fields such as psychology do not follow my preferred terminology here and may easily use the term risk for what most economists and me call ambiguity. A minority of economists also deviate from my terminology here.[2]

3.1 Uncertainty

Here, depending on context, I use S (or Ω) to denote the *state space*. Its elements are *states*, and its subsets are *events*. Usually C (or $\mathcal{C}$, or X)[3] is used to denote an *outcome set*[4], also called consequence set. *Acts* (or *prospects*), often denoted f, g, h, map states to outcomes.[5] A *simple* act takes finitely many outcomes. It can be denoted $(E_1: x_1, \ldots, E_n: x_n)$, where it is implicitly understood that $\{E_1, \ldots, E_n\}$ partition the state space

[2] By my subjective opinion, they over-extend the reach of revealed preference and endogeneity. I am amazed that some researchers who allow for exogenous information on what urns, balls, and colors are, do not want to allow for exogenous information on numbers of black and red balls being the same, do not want to consider this interesting special case, and do not want to have a special name for it.

[3] The symbol Γ is unfamiliar to many readers.

[4] In mathematical probability theory, the term outcome is often used to indicate what this paper and decision theorists call state (of nature).

[5] Following Savage (1954), we take states and outcomes as primitives and define acts in those terms. It is sometimes more natural to take acts and outcomes as primitives, and define states in their terms, mapping acts to outcomes. It may also be more natural, sometimes, to take acts and states as primitives, and derive outcomes as their product set. But Savage's setup mostly works best.

and the x_js are outcomes. This act assigns outcome x_j to each state in event E_j. Why put the events before the outcomes, and not after, in this notation? One reason is that this notation better fits with mathematical conventions of denoting functions. More importantly, it is the prevailing convention today and let us seek for uniformity of conventions. I identify constant acts with outcomes, so that preferences $\succcurlyeq$ over acts generate preferences $\succcurlyeq$ over outcomes. The last relation is so much more the same as the preceding one, rather than being different, that I use the same symbol to denote both.

Especially for finite state spaces $S = \{s_1, \dots, s_n\}$, it is sometimes convenient to take acts as n-tuples, where $x = (x_1, \dots, x_n)$ is the act assigning outcome x_i to each state s_i. Thus, also for general state spaces, the notation f_E, $f_E g$, and $\alpha_E \beta$ is as with product sets. The latter notation $\alpha_E \beta$, with event E as subscript, is concise and visual and I recommend it. Only, in contexts with heavy subscript notation it is better to write $\alpha E \beta$. As with product sets, I let subscripts refer only to states and events, and use superscripts for other indexes.

I sometimes identify one-element events $\{s\}$ with their element, the state s.

I sometimes add measure theory structure in decision under uncertainty. Then an algebra or sigma-algebra of subsets of S, denoted $\mathcal{A}$ (or Σ), is specified. Only its elements are called events, and other subsets of S are not called events. Then acts f are required to be *measurable*, i.e., $f^{-1}(I)$ is contained in $\mathcal{A}$ for every preference interval I.[6] Measure theory structure usually does not matter much for the work I do, and nonmathematicians are often not familiar with it, so I mostly omit it.

For Savage's uncertainty, EU is $\int_S U(s) dP(s)$, where I use capital P for the *probability measure* and capital U for the *utility function*. I similarly use capital U to denote the utility function in generalizations of EU. In prospect theory, I do not use Kahneman & Tversky's term value function or symbol v or V, but continue to use capital U and the term (global) utility. I often say "utility" rather than "utility function".

People often use terms "subjective expected utility" and "SEU" (for what I call EU) to emphasize that it is for uncertainty using subjective probabilities, and to emphasize that it is not risk. They then use the terms "expected utility" and "EU" only

---

[6] I is a *preference interval* if, for all $\alpha \preccurlyeq \gamma$ in I, and $\alpha \preccurlyeq \beta \preccurlyeq \gamma$, β is also contained in I.

for risk. I mostly use the same term EU for both uncertainty and risk to emphasize that they are not very different, but that the model for risk is (can be considered to be) “only” a special case of uncertainty, explained further below.

3.2 Risk

In decision under risk, C (or $\mathcal{C}$, or X) is an outcome set, as with uncertainty. *Lotteries* (or *prospects*) are probability distributions over outcomes. A *simple* lottery takes finitely many outcomes; i.e., its support is finite. It can be denoted $(p_1\colon x_1, \ldots, p_n\colon x_n)$, where it is implicitly understood that $(p_1, \ldots, p_n)$ are probabilities, i.e., nonnegative numbers that add to 1. Why the probabilities before the outcomes, and not after? For one, to be consistent with notation in uncertainty. Outome α is identified with the degenerate lottery $(1\colon \alpha)$, and $\alpha_p\beta$ denotes $(p\colon \alpha, 1-p\colon \beta)$. As with uncertainty, I recommend the concise visual $\alpha_p\beta$.

I sometimes add measure theory structure, with an algebra or sigma-algebra on the outcome set, but mostly I do not.

As mentioned before, risk is considered to be a special case of uncertainty. This assumption may not be easy to understand at first. After all, where then is the state space? But it can be seen that one can always specify an underlying state space for risk. Thus, risk is turned into the special case of uncertainty where: the state space is endowed with an objective probability measure and acts that generate the same probability distribution over outcomes are indifferent. After long experience, I recommend studying concepts as much as possible for the general context of uncertainty, and not only for the special case of risk. The general context gives better conceptual understanding.[7] People who worked too long on risk are no more able to work on uncertainty. A conclusion only for specialists: the common consequence axiom is more fundamental than the common ratio axiom.

[7] Remember the saying: “The more general the theorem, the easier its proof.”

4 Notation and terminology for rank-dependent theories, including prospect theory

I use the term rank-dependent utility both for risk, where it was introduced by Quiggin (1982), and for uncertainty, where it was introduced by Schmeidler (1989). I use the same term to emphasize that these theories are more the same than different. Quiggin's theory is "just" Schmeidler's theory for the special case of risk, where risk can indeed be considered to be a special case of uncertainty. For uncertainty many people use the term Choquet expected utility, introduced by Wakker (1990), and they use the term rank-dependent utility only for risk, but I do not follow this convention. Relatedly, I use the term prospect theory for both risk and uncertainty. Unfortunately, quite some researchers even today do not know that prospect theory, in its current 1992 version, also captures uncertainty and ambiguity, and erroneously think it is only about risk.

Avoid the terms high/low or first/last ranks, because of a linguistic ambiguity: in sports, is rank 1 higher or lower than rank 2? Say good/bad ranks.

I prefer the term prospect theory (PT) for the version of Tversky & Kahneman (1992), and the term original prospect theory (OPT)[8] for the version of Kahneman & Tversky (1979). My preference here deviates from the majority in the field, which uses terms cumulative prospect theory (CPT) and prospect theory (PT), respectively, instead. Why do I bring the inconvenience of deviating from a reasonably well-established majority terminology regarding prospect theory? Well, OPT's formula is no good and is better forgotten anyhow. Let us use the nicest term, PT, for the most important theory. I once asked Tversky, after 1992, whether the unqualified term prospect theory should be used for the old or new version, and he answered for the new version, in agreement with my preference. The term cumulative is too technical and will never appeal to broad audiences.

4.1 Rank dependence for uncertainty

In rank-dependent theories for uncertainty, we often have to rank states, or events, by the favorability of their outcomes. It then is convenient to use subscripts of the states

[8] As a confession, I always think "old prospect theory".

or events accordingly, and for instance for the act $(E_1{:}\,x_1,\ldots, E_n{:}\,x_n)$ have $x_1 \succcurlyeq \cdots \succcurlyeq x_n$, so that E_1 is the best ranked event and E_n the worst ranked.

In the preceding para, why not do the ranking the other way around, with $x_1 \preccurlyeq \cdots \preccurlyeq x_n$? Because it is very desirable for the field to have conventions of notation, and the prevailing notation is as I wrote. Relatedly, in rank-dependent weighting, for act $(E_1{:}\,x_1,\ldots, E_n{:}\,x_n)$ with $x_1 \succcurlyeq \cdots \succcurlyeq x_n$, and weighting function W, the weights of the events are $W(E_1), \ldots, W(E_1 \cup \cdots \cup E_i) - W(E_1 \cup \cdots \cup E_{i-1}), \ldots, 1 - W(E_1 \cup \cdots \cup E_{n-1})$. That is, we do top-down integration, and not bottom-up, as they are called. Why not the other way around? Because it is very, *very*, desirable, to have a uniform convention on this point, and the notation given here is prevailing today. *This notational convention is more important than many others* because without it there will be many confusions, with convex turning into concave, optimism turning into pessimism, graphs getting opposite meanings, parametric families getting different meanings, and so on. For those who do not yet follow this ranking convention: the later you retrace your steps, the higher the price you pay. ☹

In mathematical and theoretical papers, the above function W, the nonadditive set function, is often called capacity, and denoted by the Greek nu (ν) or the Roman v. In other papers, the term weighting function, put forward by Tversky, is more common, and I mostly use it to follow convention, even though it has the serious drawback that its four syllabi feel longer than the four of capacity.

Use the term rank-dependent utility (RDU), and not the term rank-dependent expected utility (RDEU). Again, brevity of terms is very important. The inventor of rank-dependent theories, Quiggin, agrees here. See:

https://personal.eur.nl/wakker/miscella/rduquiggin.txt

If, at the beginning, one did already specify subscripts in a state space $\{s_1, \ldots, s_n\}$, then one cannot use subscripts anymore to indicate rank ordering. One then would have to resort to permutations of $(1, \ldots, n)$ to indicate rank ordering. My experience is that such notation with permutations is too heavy, and is killing. My advice is to never do it. *For rank-dependent theories, do not specify subscripts of states beforehand, but reserve them to indicate rank ordering.*

For prospect theory, I assume now that the reference outcome has been specified. It is often the status quo. For monetary outcomes, it is usually taken as 0. For general outcomes, I denote it by θ. Consider a prospect $(E_1{:}\, x_1, \ldots, E_n{:}\, x_n)$, with

$$x_1 \succcurlyeq \cdots \succcurlyeq x_k \succcurlyeq \theta \succcurlyeq x_{k+1} \succcurlyeq \cdots \succcurlyeq x_n$$

Then the prospect theory value is

$$\pi_1 U(x_1) + \cdots + \pi_k U(x_k) + \pi_{k+1} U(x_{k+1}) + \cdots + \pi_n U(x_n)$$

where[9]

$\pi_1 = W^+(E_1)$, …, $\pi_i = W^+(E_1 \cup \cdots \cup E_i) - W^+(E_1 \cup \cdots \cup E_{i-1})$ for all $i \leq k$,

$\pi_n = W^-(E_n)$, …, $\pi_j = W^-\big(E_j \cup \cdots \cup E_n\big) - W^-\big(E_{j+1} \cup \cdots \cup E_n\big)$ for all $j \geq k+1$,

$U(\alpha) = u(\alpha)$ for all $\alpha \succcurlyeq \theta$,

$U(\theta) = u(\theta) = 0$,

$U(\alpha) = \lambda u(\alpha)$ for all $\alpha \preccurlyeq \theta$,

with W^+ and W^- *(event) weighting functions*, U *global utility*, u *basic utility* (utility without loss aversion), and $\lambda > 0$ *loss aversion*.

Whereas for rank dependence I insisted on top-down integration, for prospect theory I use the other integration, bottom-up, for losses. Why this seeming inconsistency? A weak argument is that it has become the prevailing convention, following Tversky & Kahneman (1992). The strong argument is that it is really best. T&K combined deep psychological insights with deep mathematical insights. They understood that their formula is most natural, both mathematically and psychologically, for reasons having to do with reference dependence that would take too much space to explain here. It later became known that T&K's dual way of integration for losses is as in the Sipos integral (Sipos 1979), an appealing alternative to the Choquet integral. But T&K discovered it independently.

Tversky & Kahneman (1992) used negative subscripts for losses, but this notation is unfortunate and is better not followed. They used the term value function instead of utility function, and symbols v (or V), but I deviate here.

4.2 Rank dependence for risk

For risk it is, again, convenient to use subscripts in agreement with ranking, with for instance $(p_1{:}\, x_1, \ldots, p_n{:}\, x_n)$ having $x_1 \succcurlyeq \cdots \succcurlyeq x_n$. In rank-dependent weighting, for

---

9 For π_1 below I assume $k > 0$, and for π_n I assume $k < n$.

lottery $(p_1{:}\,x_1,\ldots,p_n{:}\,x_n)$ with $x_1 \succcurlyeq \cdots \succcurlyeq x_n$, and weighting function w, the weights of the outcomes are $w(p_1), \ldots, w(p_1+\cdots+p_i) - w(p_1+\cdots+p_{i-1}), \ldots, 1 - w(p_1+\cdots+p_{n-1})$. As with uncertainty, we do top-down integration, and, again, it is very, *very*, desirable, to have a uniform convention on this point, and the notation given here is prevailing today. *I, again, strongly urge everyone to follow this convention of top-down integration.* I again use the term (probability) weighting function of Kahneman and Tversky. The terms probability transformation or probability distortion are sometimes used instead.

For prospect theory, consider a lottery $(p_1{:}\,x_1,\ldots,p_n{:}\,x_n)$, with reference point θ,

$$x_1 \succcurlyeq \cdots \succcurlyeq x_k \succcurlyeq \theta \succcurlyeq x_{k+1} \succcurlyeq \cdots \succcurlyeq x_n$$

Then the prospect theory value is

$$\pi_1 U(x_1) + \cdots + \pi_k U(x_k) + \pi_{k+1} U(x_{k+1}) + \cdots + \pi_n U(x_n)$$

where[10] $\pi_1 = w^+(p_1), \ldots, \pi_i = w^+(p_1+\cdots+p_i) - w^+(p_1+\cdots+p_{i-1})$ for all $i \leq k$,

$\pi_n = w^-(p_n), \ldots, \pi_j = w^-\big(p_j+\cdots+p_n\big) - w^-\big(p_{j+1}+\cdots+p_n\big)$ for all $k+1 \leq j$,

$U(\alpha) = u(\alpha)$ for all $\alpha \succcurlyeq \theta$,

$U(\theta) = u(\alpha) = 0$,

$U(\alpha) = \lambda u(\alpha)$ for all $\alpha \preccurlyeq \theta$,

with w^+ and w^- probability weighting functions, U global utility, u basic utility, and λ loss aversion. For discussion, see uncertainty.

5 Miscellaneous

Wang (2024) provided many suggestions for writing mathematics.

I use "if", and not "if and only if", in definitions, and italicize the concept to be defined. Thus, the preference relation $\succcurlyeq$ is *complete* if $x \succcurlyeq y$ or $y \succcurlyeq x$ for all prospects x, y, and it is a *weak order* if it is complete and transitive. I just defined completeness and weak ordering here. It is important to distinguish definitions from regular statements. I regret that many authors do not follow an explicit convention for indicating definitions. They themselves know what are definitions and what are

[10] For π_1 below I assume $k > 0$, and for π_n I assume $k < n$.

statements, but their readers can only guess. As written, my convention is to use italics to indicate definitions.

Relatedly, organize definitions and notation to easily find back. Put no assumptions unorganized in the flow of the text. If, to understand Theorem 6, I have to read the entire preceding text line by line to find out if there was an extra assumption stated somewhere randomly in the flow of the text: no good. General notation and terminology is listed at the beginning of the first formal section. I usually summarize assumptions in a displayed "Structural Assumption" that I refer to in theorems. This makes it very easy for readers to understand theorems by minimizing the required texts to check out or memorize.

Use abbreviations sparingly. Besides standard abbreviations such as EU, introduce no more than two or three new abbreviations.

I use the terms nondecreasing and strictly increasing, and not the term increasing because for the latter there is, unfortunately, no standard convention. Nonincreasing and strictly decreasing are similar.

For nonmathematicians, *linearity* of a function f often means what mathematicians call *affinity*: $f(\lambda\alpha + (1-\lambda)\beta) = \lambda f(\alpha) + (1-\lambda)f(\beta)$ for all $0 \leq \lambda \leq 1$. For mathematicians, *linearity* further requires that f assign value 0 to 0 (or to the origin).

I never use the symbols $\subseteq$ or $\supseteq$, but only $\subset, \supset$. I may thus write $A \subset A$.

To denote the set of reals, I prefer $\mathbb{R}$ to $\mathfrak{R}$, but this may be a minority view. $\mathbb{R}^+ = [0,\infty)$, and $\mathbb{R}^{++} = (0,\infty)$.

$\mathbb{N}$, the set of natural numbers, denotes $\{1,2,\ldots\}$, and does not include 0. $\mathbb{N}^0 = \{0,1,2 \ldots\}$.

Note the choice of $\cdots$ versus ..., depending on the surrounding symbols, as in

$$= \cdots =$$

versus

$$, \ldots, .$$

I do start sentences with symbols if the latter are capitals. Sometimes even if no capitals, although many people dislike that.

Other things equal, preference orders between outcomes are $\delta \succcurlyeq \gamma \succcurlyeq \beta \succcurlyeq \alpha$ where I take in mind that γ is good and β is bad. If more symbols are needed then I add δ thinking delicious and/or α thinking awful. Such lifelong habits make life easier. More importantly, in general, choose notation, symbols, and terms to help memory. Let mathematical symbols correspond with first letters of the concepts they refer to. Bad examples are Kaheman's widely used but unfortunate terms "system 1" and "system 2" and Sen's unfortunate terms "property α" and "property β".

Write $1, \ldots, n$ and not $1,2, \ldots, n$.

I would much like to use logical symbols such as $\forall, \exists, \Rightarrow, \Leftarrow, \Leftrightarrow$, because they are concise and visual, but, unfortunately, most economists do not know them. I may happily use them for theoretical journals such as Theoretical Economics, Journal of Economic Theory, and Economic Theory, but even for Econometrica I do not use them.

Uncommon Greek letters are not known by many non-Westerners and are better used sparingly. Well-known are the Greek letters $\alpha, \beta, \gamma, \delta, \varepsilon, \mu, \lambda, \sigma, \tau$, and next come θ, ν, ω.

I recommend one continuous numbering of theorems, lemmas, definitions, and so on. So, Definition 1, Lemma 2, Theorem 3, Definition 4, Theorem 5, Observation 6, etc. The big pro is that it then is easier to find back cross-references to such items. Some people have a linguistic desire that, if there is a Theorem 5, there should also be a Theorem 4, but I assign utility 0 to this desire. Just drop it! My convention does make it harder to know how many theorems a paper has, but this drawback does not outweigh the aforementioned pro.

I recommend numbering most or even all displayed equations, also those not cited in the paper. Other authors of other papers may want to cite such equations. I cannot think of a drawback of my convention.

Mathematical expressions can be treated as atomic in sentences, but can also be taken to provide grammatical inputs such as verbs. Thus, one can write "We show $x = y$" but also "We show that $x = y$".

6 Conclusion

Conciseness of notation, and uniform conventions, are desirable. I hope that this note can serve these goals.

ACKNOWLEDGMENT. Han Bleichrodt made useful comments.

References

Kahneman, Daniel & Amos Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, 263–291. https://doi.org/10.2307/1914185

Quiggin, John (1982) "A Theory of Anticipated Utility," *Journal of Economic Behaviour and Organization* 3, 323–343. https://doi.org/10.1016/0167-2681(82)90008-7

Savage, Leonard J. (1954) "*The Foundations of Statistics.*" Wiley, New York. (2nd edn. 1972, Dover Publications, New York.)

Schmeidler, David (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571–587. https://doi.org/10.2307/1911053

Šipoš, Ján (1979) "Integral with Respect to a Pre-Measure," *Mathematica Slovaca* 29, 141–155. https://doi.org/10.2307/1911053

Tversky, Amos & Daniel Kahneman (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323. https://doi.org/10.1007/BF00122574

Wakker, Peter P. (1990) "Under Stochastic Dominance Choquet-Expected Utility and Anticipated Utility are Identical," *Theory and Decision* 29, 119–132. https://doi.org/10.1007/BF00126589

Wang, Ruodu (2024) "Suggestions for Writing Mathematics in Scientific Papers." https://sas.uwaterloo.ca/~wang/files/writing.pdf