

Theory Choice and the Intransitivity of 'Is a Better Theory Than'*

Peter Baumann^{†‡}

There is a very plausible transitivity principle for theory choice. It says that if all criteria of theory evaluation are considered, and theory A is a better theory than theory B, and theory B is a better theory than theory C, then A is a better theory than C. I argue against this principle. It turns out that whenever there are two or more relevant and independent criteria of theory evaluation, and that whenever at least of one the criteria is 'nonlinear' in a certain sense, there may be violations of transitivity that do not violate any standards of rationality (of theory choice). This shows, again, that theory choice cannot be seen as merely the application of given rules of rational theory choice.

1. Introduction. Suppose there is good evidence for assuming that Mary has gone swimming (p): She has told me so yesterday. However, there is even better evidence that she has not gone swimming, but is in her office instead (q): The office door is wide open—as I can clearly see. But there is even better evidence that she is not in her office, but in the copying room (r): I have just met Jack, who never ever lies and who is very reliable; he tells me that he's just seen Mary at the copying machine.

Let us also assume that the quality of the evidence is the only thing that counts with respect to the rational acceptance of the three propositions in the example. We can then say that if the person has better evidence for one proposition (q) than another (p), then she has better reasons to accept q than to accept p . One could also say: She is better justified in accepting q than in accepting p . Let ' $R(q, p)$ ' stand for 'the person [e.g.,

*Received December 2003; revised July 2004.

†To contact the author, please write to: Department of Philosophy, University of Aberdeen, Old Brewery, Aberdeen, United Kingdom, AB24 3UB; e-mail: p.baumann@abdn.ac.uk.

‡I would like to thank the following persons for comments and discussions: Gisela Cramer, Hugh Lacey, Abdul Raffert, Howard Sankey, and two anonymous referees.

Philosophy of Science, 72 (January 2005) pp. 231–240. 0031-8248/2005/7201-0013\$10.00
Copyright 2005 by the Philosophy of Science Association. All rights reserved.

me in the example above] has better reasons, all things considered, to accept q than to accept p .⁷ In our example the following is true:

$$R(q, p) \ \& \ R(r, q).$$

There seems no doubt whatsoever that this implies $R(r, p)$. Our example thus suggests a general principle about reasons (or justifications) for accepting propositions (or sets of propositions, or theories) as true:

(T) For all propositions (or sets of propositions, or theories) p , q , and r :

$$[R(q, p) \ \& \ R(r, q)] \rightarrow R(r, p).$$

One can call this the principle of the ‘transitivity of reasons’.¹

(T) seems to work in everyday cases, as well in cases of theory choice. In the latter case, R is a relation that holds between two theories A and B ($R(A, B)$), if and only if A is (overall) a better theory than B—taking all relevant criteria of theory evaluation into account. The relevance of those criteria is, of course, highly controversial; however, the argument in this paper does not presuppose any particular view on relevance; we can thus leave this topic aside here. The argument that follows is so general in nature that it holds (if, indeed, it does hold) no matter what one’s particular views about rational acceptability of theories (R) are. We can thus also leave this topic aside here.

Suppose there are three competing theories under discussion in a scientific community. Let us assume for a short moment that the quality of the evidence that speaks in favor of a theory is the only thing that matters when it comes to rational acceptability of theories. After some time, the scientists come to the conclusion that theory 1 is better supported by the evidence than theory 2, and that theory 2 is better supported by the evidence than theory 3. It seems that there is no further discussion necessary about which theory to accept or choose: Theory 1 must also be better supported by the evidence than theory 3. Hence, it must be the best theory in this situation (given, of course, that only the quality of the evidence counts).

(T) looks like an essential aspect of the rational acceptance of the rational belief in propositions, sets of propositions, or theories. Apart from all that, (T) looks like an (almost) trivially true statement about the nature of rational epistemic choice. But, it is not. I will argue here that it is neither trivial nor true. This goes against a basic assumption of

1. I am only talking about epistemic, not pragmatic, reasons to accept propositions or theories.

decision theoretic approaches to epistemic rationality.² In the following, I will restrict myself to the case of theory choice.³ I cannot present a case study here (this would require another paper), but instead, I only intend to present a general argument. It will proceed in two steps as detailed in Sections 2 and 3.

2. First Step: The Basic Idea. There is—let us assume—more than one criterion for the evaluation of scientific theories: support by empirical evidence counts a lot, but so does explanatory power, simplicity, and other factors.⁴ Assume further, that the plurality of criteria is irreducible such that none of them can be explained in terms of one or more of the other criteria. Let us, for the sake of argument, also assume that the three criteria just mentioned are all the relevant criteria, and that they are independent in the sense that the 'quality' of a theory in one dimension does not imply anything (or at least not very much) about how good the theory is with respect to the other dimensions of evaluation.⁵ Suppose that three theories—call them 'A', 'B', and 'C'—are under discussion in a scientific community. Let us further suppose that the following is also true. A can clearly explain more than C, and it can explain everything C explains, but in a much more convincing way. On the other hand, C is clearly better supported by the data and not as complicated (a theory) as A. Which of the two theories should one accept?

2. Epistemic choice, or theory choice, is just a special case of choice in general. There are thus parallels between this paper and the (recent) discussions about transitivity in decision theory. For the 'classical' view that rational choice requires transitivity cf., e.g., Ramsey 1990, 75, 78; von Neumann and Morgenstern 1953, 26–27; Luce and Raiffa 1957, 23, 25, 28; Savage 1972, 18; Jeffrey 1974, 391. Cf. for different views: Tversky 1969; Schumm 1987; Bar-Hillel and Margalit 1988; Fishburn 1991; Anand 1993, 55ff.; Gendin 1996. Cf. also Broome 1993, 53ff., Broome 1991, 80ff., and the discussion between Hughes 1980, Lee 1984, Philips 1989, and Rawling 1990. For empirical studies on intransitive preferences, cf. Tversky 1969.

3. I will not consider the case of theories that are so poor that they are not worth serious consideration by a theorist or scientist. The kinds of theories I have in mind would all pass a threshold of minimal acceptability. They are all 'good' in this sense. What follows is about problems with the alleged transitivity of the *better than* relation among those (good enough) theories.

4. Cf., e.g., Kuhn 1977, 321–322; Lacey 1999, 58ff. Bayesians (cf. Howson and Urbach 1989) would deny this; I cannot go into this here. I am also assuming that none of the criteria is mistaken in one way or another. If one has doubts concerning one of the criteria above, one can easily choose alternative ones (given pluralism of criteria) such that the structure of the argument would remain unchanged.

5. We could also assume that they are mostly independent, but for the sake of simplicity, it is more convenient to assume total independence; nothing substantial hinges on this simplification.

A lot depends on how one weighs the different criteria against each other: How important, for instance, is explanatory power in comparison with support by data? One rather extreme view would have it that there is a lexical ordering of different criteria. This seems pretty unrealistic, though. Another extreme view would say that there is no way to compare the relative importance of the different evaluative criteria relevant for theory choice; this would lead to some kind of incommensurability.⁶ However, we can resist this view. There might be no context-independent and general ways of weighing the different criteria against each other, but in at least many *particular situations* of scientific theory choice, scientists seem able to clearly weigh the criteria against each other (cf., e.g., Sankey 1994, 176). Only for the sake of simplicity, will I assume for a moment that in our example that the three criteria are equally important (nothing essential hinges on this assumption; I will later skip this assumption).⁷ All this suggests is that the scientists have better reasons overall to prefer C to A. This might change over time, as new data come in, or more sophisticated versions of the original theories are developed; in the particular situation at hand, however, C seems to be the ‘better theory’ (as we can also say).⁸

Let us now assume that the following is true about theories A, B, and C. With respect to explanatory power, A is better than B, and B is better than C. With respect to support by data, B is better than C, and C is better than A. With respect to simplicity, C is better than A, and A is better than B. We can illustrate this in the following way (the first, second, and third row gives us the best, second best, and worst theory, as evaluated by the relevant criterion):

Theory Ranking	Explanatory Power	Support by Data	Simplicity
Best	A	B	C
Second best	B	C	A
Worst	C	A	B

Given our assumptions, we would have to say that all things considered,

6. Though not of the Kuhnian sort (cf. Kuhn 1962). For an argument of the above kind, cf. Sankey 1995.

7. A similar argument can be made for the case in which the criteria are of different importance. The exposition would just become a bit more complicated.

8. For purely stylistic reasons, I will not explicitly mention the reference to a particular context or situation from now on.

C is a better theory than A (it 'beats' A 2:1 on the relevant criteria). Also, A is better than B, and B is better than C. In other words:

$$R(C, A) \ \& \ R(A, B) \ \& \ R(B, C).$$

This, however, clearly goes against (T). Since it is hard to see why there should be something wrong with the individual evaluations, it seems to follow that the relation *is a better theory than* is not transitive (with 'better' understood as 'better all things considered'). There is no best theory in this context. Whichever theory one chooses, there always seems to be a better theory on offer. I have assumed pluralism in the sense that there is an irreducible plurality of criteria without lexical ordering. If one wants to keep the question of pluralism open, one would have to replace the thesis above by a weaker, conditional thesis: If pluralism is true, then rational theory choice does not obey the transitivity rule.

3. Second Step: Further Aspects and Some Objections. One could object that the above comparison of the three theories neglects the possibility that, e.g., the difference in explanatory power between A and C might be much 'bigger' or should weigh much more than their difference with respect to support by empirical data. Should one not take such differences into account? It is not clear whether this makes sense at all: Can we, for instance, measure different degrees of explanatory power of scientific theories, and compare it with differences in simplicity? Let us, for the sake of the argument (or, for the sake of the objection) assume that we can (more or less roughly) measure different degrees to which a theory can meet a given criterion. Let us further—and also for the sake of the argument—assume that there are just two relevant criteria: explanatory power and simplicity; we can assume that our three theories do not differ with respect to how well they are supported by evidence. Would transitivity of rational choice always hold under these conditions? I will argue that the answer is 'no'.

Suppose there are three competing theories: D, E, and F. D explains things much better than E, but is also more complicated. However, the difference in simplicity is not so big that it really counts much, whereas the difference in explanatory power really matters. Hence, all things considered, D is better than E. Similar things hold for E and F: E is a bit more complicated, but explains much better than F. Hence, E is better than F. What about D and F? Suppose F is clearly and remarkably simpler than D—so much simpler that it really matters. It matters more than the difference in explanatory power. Hence, F is better than D, all things considered. Again, we have an intransitive ranking among our three theories. There is no best theory, and for any given theory from our three competing theories, there is always a better theory. Again, there is no

reason to believe that there is something wrong with the different evaluations of the theory.⁹ Or, to put it in terms of our weaker, conditional thesis: If pluralism is true, then there is rational acceptability of intransitive rankings of theories.

One might want to propose the following way out of intransitive circles. In a first step, one determines for each of the different criteria how ‘good’ a given theory is with respect to that criterion. One could do this in a quantitative way and assign real numbers that measure the ‘goodness’ of a theory in one dimension (defined by one of the criteria of evaluation).¹⁰ Assuming that the different criteria constitute different dimensions of one and the same quality of ‘goodness’ of a theory, one can then compute the overall ‘goodness’ of a theory with respect to the different criteria. Since one can do all this for each of the theories under review one can in the end compare the (absolute) ‘goodness’ of different theories in quantitative terms.

There are many problems with this proposal. One is quite simple: It assumes monism, that is, the idea that in the end there is only one criterion that matters, namely ‘goodness’. Apart from the doubts concerning monism mentioned above, there is an additional problem with this proposal: It is not clear at all what ‘goodness’ is supposed to be (if it does not reduce to the different aspects of evaluation captured by the different criteria). It does not seem to be an independent, substantial criterion at all. Whenever we say that a given theory is a good one, we have to specify criteria of evaluation with respect to what makes it a good one. But, then the alleged way out of intransitivity leads us back to where we started.

Another way to avoid intransitivity of rational theory choice would be to introduce some kind of ‘metaevaluation’ of the criteria of theory choice. I have assumed here that no criteria are mistaken in any way. Metaevaluation could then only be about the relative importance of the different criteria. However (as should be plausible by now), the argument presented here does not only not exclude or ignore this aspect, but rather takes it into account from the start (see above). Larry Laudan (cf. 1987, 1990) defends a ‘normative naturalism’ according to which we can resolve deep methodological disputes by engaging in some metaevaluation of the different methods used. According to Laudan, one method is better than another one, if it does a better job in reaching its cognitive aim. Be it as it may—Laudan’s point concerns the evaluation of methods of inquiry, and not of criteria of evaluation of theories (which is something different).

9. One would have to beg the question in favor of transitivity in order to hold this view.

10. Perhaps using methods along the lines of those proposed by Ramsey 1990, 68–78 and by von Neumann and Morgenstern 1953, 17–19.

Hence, it does not seem to make sense to extend Laudan's metamethodology to the evaluation of criteria. Apart from that: What constitutes a valuable cognitive aim depends on the correct criteria of evaluation. Laudan's metamethodology takes the latter as given—which we cannot do here if we want to engage in the proposed metaevaluation of criteria.¹¹

What now could explain the (rational) possibility of intransitive rankings of theories? Consider our initial example about Mary: It does not speak against (T). We can now see why: There was only one criterion of evaluation. In the case of theories D, E, and F, however, there are two independent criteria: simplicity and explanatory power. Furthermore, at least one of these criteria (simplicity) is 'nonlinear' in the sense that small differences (say between theories F and E, or E and D above) do not matter much (or not at all) to the chooser, but as they add up (see the case of theories F and D above), their importance for theory choice might suddenly be much bigger.¹² There might be thresholds of complexity: Small differences of complexity might be completely negligible up to a certain point; if they are even slightly greater than the threshold value, then they might suddenly make all the difference (or at least a huge difference). But even without such thresholds, the function that maps degrees of complexity onto degrees of 'goodness' of a theory, with respect to simplicity, might not be linear.¹³

In other words: If there are two or more relevant criteria for the evaluation of theories, and if at least one of those criteria is nonlinear, then

11. For the limits of metaevaluation, cf. Sankey 1994, 172–174 and Brown 1988.

12. What exactly does 'nonlinearity' mean? Suppose there are two relatively unimportant phenomena *P* and *Q*. Suppose further that the only relevant difference between two given theories T1 and T2 is that T1 can explain *P* but not *Q*, whereas T2 can explain *Q* but not *P*. Suppose also that the theorist is indifferent between T1 and T2. In other words, it is as (un)important to be able to explain *P* as to explain *Q*. We can now use this measure of equal (un)importance for an explanation of nonlinearity. Take theories F and E above and assume, again, that F is only slightly simpler than E: The theorist is indifferent between the advantage in simplicity F has over E (and E over D) on the one hand, and the explanatory advantage of being able to explain *P* over not being able to explain *P* on the other hand. Now, if the chooser is not indifferent between (a) the advantage in simplicity F has over D and (b) the explanatory advantage of being able to explain *P* and *Q* over not being able to explain either *P* or *Q*, then the simplicity criterion is not linear (at least when it comes to differences as those between F, E, and D). More could be said about nonlinearity but these remarks should be sufficient here.

13. Again, this discussion is based on the controversial assumption that there are degrees of complexity, evidential support, explanatory power, etc., that we can somehow measure. One can deny all this. This, however, does not amount to an objection to the argument above, but rather to a rejection of the whole discussion: One would neither accept (T). I can live with that, because my aim here is to throw doubt on (T).

transitivity of rational theory choice (i.e., (T)) might be violated. Multiplicity of criteria is necessary, but not sufficient, for the rational acceptability of some intransitive preferences among theories. The same holds for nonlinearity: It is necessary for it, but not sufficient. Pluralism and nonlinearity are individually necessary, and jointly sufficient conditions, for the rational acceptability of some intransitive preferences among theories. I am not saying that rational theory choice will always violate transitivity (according to (T)) if those two conditions are given; the point is, rather, that it can (but need not) be rationally acceptable to violate transitivity, given those conditions. It is very realistic to assume that there is more than just one independent criterion for the evaluation of scientific theories and that there is no lexical ordering of criteria. It is also realistic to assume that at least some of those criteria would be nonlinear. Hence, we have good reasons to doubt and reject (T). Or, if one wants to keep the question of pluralism open: If pluralism is true, then rational theory choice does not obey the transitivity axiom (T).

The upshot is this: The violation of transitivity, as explained by (T), might be a common phenomenon in science—whenever there is a choice between more than two competing theories. More important even: Such transitivity might be violated in such a way that there is no reason at all to say that there is something wrong or irrational about the way the scientists evaluate the different theories. The life of science could just be like that. In some cases (not all, of course), rational choice between theories violates the transitivity assumption. And, similar things can presumably be said about cases from everyday life, at least if they show a certain complexity and are not as simple as the initial example about Mary.¹⁴

One final objection: Have I not completely ignored the normative question of how someone who chooses among several theories ought to choose? Should I not say something about how a rational chooser ought to weigh the different criteria of theory choice? Isn't there, for instance, a normative question about whether a given criterion ought to be linear or not (and if not: in what ways it ought to be nonlinear)? I have indeed not dealt with this whole normative dimension. Fortunately, this does not do any damage: Take any normative view about theory choice, and the same problem with intransitivity will come up within that framework (given criteria pluralism). Hence, we can leave the whole normative dimension aside here.

14. As I have said above, I have restricted myself to minimally acceptable theories. The argument here does thus not exploit features of those theories that are unworthy of consideration.

4. Conclusion. But how is rational comparison between theories or between propositions possible if all that is true? If the possibility of a rational answer to the question, 'Which theory should we choose?' presupposes that there is always a best theory (or a group of equally good theories such that no theory is better), then there cannot always be a 'rational' answer in that sense. However, if what I have said above is true, then we should not be too obsessed with the idea that there must always be a best theory. What matters is that the theories we choose from are good enough. Furthermore, the phenomenon of intransitivity might be a passing one: As new additional evidence comes up—to mention just one dimension of evaluation—the intransitivity might disappear. Intransitivity might be an unstable phenomenon; it is hard to think of a case in the history of science in which intransitivity did not disappear after a relatively short time.

All this does not lead to Kuhnian incommensurability. It could still be true that for all theories (or, alternatively, propositions) A and B:

$$R(A, B) \vee R(B, A) \vee E(A, B)$$

where *E* stands for 'is equally good a theory as'. But in another respect, one would probably be tempted to agree with Kuhn: Theory choice—or, more generally, the acceptance of certain propositions—is not (always) a purely rational process, or just the application of some algorithm. It also includes more 'pragmatic' elements (not just epistemic criteria): what appeals to a group of researchers at a time, and what they can manage to persuade each other of.

REFERENCES

- Anand, Paul (1993), *Foundations of Rational Choice Under Risk*. Oxford: Clarendon.
- Bar-Hillel, Maya, and Avishai Margalit (1988), "How Vicious Are Cycles of Intransitive Choice?", *Theory and Decision* 24: 119–145.
- Broome, John (1991), "Rationality and the Sure-Thing Principle", in Gay Meeks (ed.), *Thoughtful Economic Man*. Cambridge: Cambridge University Press, 74–102.
- (1993), "Can a Humean Be Moderate?", in Raymond Gillespie Frey and Christopher W. Morris (eds.), *Value, Welfare, and Morality*. Cambridge: Cambridge University Press, 51–73.
- Brown, Harold I. (1988), *Rationality*. London: Routledge.
- Fishburn, Peter C. (1991), "Non-Transitive Preferences in Decision Theory", *Journal of Risk and Uncertainty* 4: 113–134.
- Gendin, Sidney (1996), "Why Preference Is Not Transitive", *The Philosophical Quarterly* 46: 482–488.
- Howson, Colin, and Peter Urbach (1989), *Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court.
- Hughes, R. I. G. (1980), "Rationality and Intransitive Preferences", *Analysis* 40: 132–134.
- Jeffrey, Richard C. (1974), "Preference among Preferences", *The Journal of Philosophy* 71: 377–391.
- Kuhn, Thomas S. (1962), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- (1977), "Objectivity, Value Judgment, and Theory Choice", in Thomas S. Kuhn,

- The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press, 320–339.
- Lacey, Hugh (1999), *Is Science Value Free? Values and Scientific Understanding*. London: Routledge.
- Laudan, Larry (1987), “Progress or Rationality? The Prospects for Normative Naturalism”, *American Philosophical Quarterly* 24: 19–31.
- (1990), “Normative Naturalism”, *Philosophy of Science* 57: 44–59.
- Lee, Richard (1984), “Preference and Transitivity”, *Analysis* 44: 129–134.
- Luce, R. Duncan, and Howard Raiffa (1957), *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.
- Philips, Michael (1989), “Must Rational Preferences Be Transitive?”, *The Philosophical Quarterly* 39: 477–483.
- Ramsey, Frank Plumpton (1990), “Truth and Probability”, in David Hugh Mellor (ed.), *Philosophical Papers*. Cambridge: Cambridge University Press, 52–94.
- Rawling, Piers (1990), “The Ranking of Preference”, *The Philosophical Quarterly* 40: 495–501.
- Sankey, Howard (1994), “Judgment and Rational Theory Choice”, *Methodology and Science* 27: 167–182.
- (1995), “The Problem of Rational Theory Choice”, *Epistemologia* 18: 299–312.
- Savage, Leonard J. (1972), *The Foundations of Statistics*. New York: Dover Publications.
- Schumm, George F. (1987), “Transitivity, Preference and Indifference”, *Philosophical Studies* 52: 435–437.
- Tversky, Amos (1969), “Intransitivity of Preferences”, *Psychological Review* 76: 31–48.
- von Neumann, John, and Oskar Morgenstern (1953), *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.