

Supplying Planks for Neurath's Boat: Can Economists Meet the Demands of the Dynamics of Scientific Theories?

Hans Rott, Regensburg

1. Introduction

According to Otto Neurath, the conduct of science consists in a large undertaking of setting up and maintaining systems of statements:

In unified science we try ... to create a *consistent system* of protocol statements and non-protocol statements (including laws). When a new statement is presented to us we compare it with the system at our disposal and check whether the new statement is in contradiction with the system or not. If the new statement is in contradiction with the system, we can discard this statement as unusable ('false'), for example, the statement: 'In Africa lions sing only in major chords'; however, one can also 'accept' the statement and change the system accordingly so that it remains consistent if this statement is added. The statement may then be called 'true'.

The fate of being discarded may befall even a protocol statement. There is no '*noli me tangere*' for any statement though Carnap claims it for protocol statements (Neurath 1932/33, pp. 94-95)

Neurath's coherentist epistemology – famously expounded in the *Protokollsatze* with Moritz Schlick – rejects any commitment to a distinguished set of sentences that might be used as firm foundations upon which to build the whole of science. The scientist's task is to keep his system coherent, not to pay undue respect to a predetermined class of privileged truths. Neurath was very clear about the fact that there are many consistent systems that might be offered as accounting for the observed phenomena. In contrast to Schlick, Neurath indeed welcomed the idea that there is no unique "criterion of truth". Scientists have to make decisions which of these systems to adopt, and indeed there is no escape from the making of decisions in science.

Neurath, however, said disappointingly little about *how* to choose the most adequate system of propositions among multiple coherent candidate systems. In "Sociology in the Framework of Physicalism" (1931, p. 66), he remarked that when encountering some recalcitrant data, it is hard to decide in favour of the data and a revision of the system one has been using so far. A little later, in "Radical Physicalism and the 'Real World'" (1934, p. 109), he held that scientists should aim at making their theories agree with as many protocol statements as possible.¹ While Neurath took the latter requirement to be a mark of empiricism, we can also interpret it as an element of

¹ Neurath often emphasised that constructing the set of protocol statements is a matter of choice, see e.g.: "Without an adequate selection [of observation statements] at the start, fertile scientific work is impossible." (1931, p. 84)

conservatism: When trying to accommodate new observations, scientists should never sacrifice the observations that had been successfully accounted for before the new observation was made. This fits together with the general reluctance ("*wozu man sich im allgemeinen schwer entschließt*") of scientists to revise a corroborated system that Neurath was speaking about in the earlier paper.

The textual evidence for this interpretation, however, is thin. It was only Neurath's opponent in this discussion, Schlick, who in his paper "The Foundation of Knowledge" made a clear suggestion as to how the problem of choice involved might be solved. In a way congenial with Neurath, he proposed to employ a *principle of economy*:

By what mark, then, are we to distinguish these statements which themselves remain unaltered, while all others must be brought into agreement with them? ... The most obvious recourse would doubtless be to find the rule for which we are searching in some kind of *economy principle*, namely, to say: we are to choose those as basic statements whose retention requires a *minimum* of alteration in the whole system of statements in order to rid it of all contradictions. ...

It is worth noticing that such an economy principle would not enable us to pick out certain statements as being basic once and for all, ... This would, of course, no longer be the pure *coherence* viewpoint, but one based on *economy* (Schlick 1934, pp. 216–217)

Schlick thought that the proponents of the coherence theory of scientific knowledge he attacked "did in fact take the economy principle as their guiding light, whether explicitly or implicitly", but he argues, against them, that the epistemic merits of a statement consist in its pedigree rather than its pragmatic utility.²

The present paper takes a look at what may be viewed as a successor project to Schlick's idea of supplementing Neurath's picture of the structure and dynamics of science by an economy principle. In the 1980s, logical models were developed that explicitly addressed the problem of maintaining consistency in systems of statements, and it was often proclaimed that the particular way of choosing among consistent successor theories was guided by an economy principle very much like Schlick's.³ I am referring to the *theory of theory change*, which is also known under the name *belief revision*. The theory has been rather limited, because it postulates that new information be always accepted. In this respect the model falls behind Neurath's (1932/33, pp. 94–95) insight that in case

² Neurath himself rejected the term "coherence theory" as a label for his point of view, but in his reply to Schlick (Neurath 1934), he does not even begin to comment on Schlick's (re-)interpretation of his account as based on the economy principle. One might speculate whether there is a path leading from Richard Avenarius' *Principle of Least Action* or Ernst Mach's *Principle of the Economy of Thought* to Schlick's mentioning of an economy principle here. Mach was, to be sure, one of the most influential scientists at the turn of the century, and Neurath's founding of the *Verein Ernst Mach* in 1928 was one of the major steps in the development of the Vienna Circle. But Mach's principle was much closer in spirit to Occam's razor than to the economy principle formulated by Schlick.

³ The landmark paper is Alchourrón, Gärdenfors and Makinson (1985), for book-length treatments see Gärdenfors (1988), Hansson (1999) and Rott (2001).

of conflict one might *either* revise the theory *or* reject the new statement as "useless" or "false".⁴ However, this seems to me a minor defect, one that does not invalidate the discussion below and that has recently been rectified anyway.⁵

Where does this account leave deduction and induction in the sciences? According to Neurath, inductive reasoning always involves decisions as to which sentences to adopt, while deductive logic serves as a standard against which to measure whether potential theories are free of contradictions or not. *Logic in a narrow sense* refers to formal models of deductive reasoning, axiomatic systems which should preferably come together with some semantic underpinning. *Logic in a wider sense* is the theory of good reasoning, providing us not only with the means to check candidate theories for consistency but also with a methodology for deciding which of the candidate theories to adopt. Neurath (1934, p. 106) rightly laid stress on the fact that theory choice is based on *extralogical* considerations, but in this he clearly referred to the narrow conception of logic. I will in this paper proceed on the assumption that logic does include rules that are relevant for processes of belief formation and theory choice (a kind of inductive reasoning, see Spohn 2003). We shall see that we can clearly distinguish an idea of *economical* behaviour (or *economy*) from behaviour as recommended by *economics* (*economic* behaviour). We will ask to what extent the two ideas have as a matter of historical fact become embodied in formal models for theory formation and transformation, and we will address the question to what extent they *should* be respected in these models. In doing so, I will hardly say anything about the role logic has to play in (empirical) science, real or reconstructed. This paper is not a piece of formalised philosophy of science. Rather, I want to have a look at the converse question about the role that the science of economics may play in logic broadly conceived. In order to do this, we first have to get an idea of what economics and economy are.

2. What is economics?

One way of finding out what a term means is to look at the science that is supposed to study it. For the word 'economic', this seems to be an easy task. We just have to look at what economics is about. Dictionaries define 'economics' as the scientific study of the production, distribution and consumption of goods, services and wealth, or more concretely, the study of the system of trade, industry, money etc. But we want to dig deeper. According to Francis Y. Edgeworth (1881, p. 16), "[t]he first principle of Economics is that every agent is actuated only by self-interest." This

⁴ This insight was termed *the Neurath Principle* by Haller (1982a,1982b).

restriction to a completely selfish attitude has long been removed. The formal part of Edgeworth's idea, however, remains valid. In the words of Herbert Simon: "The rational man of economics is a maximiser, who will settle for nothing less than the best."⁶ This still seems to be the dominant view, as is shown by a few more recent statements. Hausman (1998) gives the following summary of the dominant school in the 20th century: "The main 'orthodox', 'neoclassical', or 'neo-Walrasian' school models economic outcomes as equilibria in which individuals have done as well for themselves as they could given their preferences and the constraints on their choices." In another handbook article, Rosenberg (1995) tells us about the "assumptions of the 'economic man': that all agents have complete and transitive cardinal or ordinal utility rankings or preference orders and that they always choose that available option which maximises their utility or preferences". Earlier the same author characterised economics as guided (or misguided) by an "extremal intentional research programme" (Rosenberg 1983).⁷ And what is being maximised is utility. According to Broome (1999, p. 21-22): "'utility' acquired the meaning: *the value of a function that represents a person's preferences*. ... The first principle of economics is ... utility theory, ... modern, axiomatic utility theory ..."

What is common to these very abstract formulations of the basic tenets of modern economics is that economic agents are viewed as having definite preferences, and that when choosing actions or commodities, they aim at satisfying their preferences as well as the circumstances allow. It is important to note that not just any kind of preference is considered to be appropriate. Preferences have to be transitive and complete (technically speaking, they must be "preorderings" or "weak orderings"), in order to be representable by a suitably chosen utility function. Ties in the preferences are permitted, of course, but incomparabilities are ruled out. Put very pointedly, one can say that economics is based on (and, in a very abstract sense: is an elaboration of) the theory of rational choice. Here, *choice* is called *rational* or *coherent* if and only if it is representable by a preference relation, and a *preference relation* in turn is considered to be *rational* if and only if it is representable by a *utility function*.⁸

⁵ See the papers on "non-prioritised belief revision" collected in Hansson (1997).

⁶ The quotation is taken from McFadden (1999, p. 73). Simon himself famously advocated an alternative model of rationality, viz., that of *satisficing* instead of maximising. According to this model, economic agents set out for achieving a certain level of aspiration, without investing any extra efforts to find out whether they could obtain something even more valuable.

⁷ According to Rosenberg (1983), economics is not an empirical science, but a branch of applied mathematics – but he does not say applied to *what*. My picture below is more in line with the view that economics is a normative discipline, a theory of rationality.

⁸ It takes more than transitivity and completeness for a preference relation to be representable by a utility function. An additional technical continuity property is needed (Debreu 1959, 54-59). A concise presentation of ordinal and expected utility theory is given by Hausman and McPherson (1996, Chapter 3: "Rationality").

Economics, then, is about *rational* rather than the *actual* behaviour of individuals. The laws of economics should not be expected to be empirically adequate, they are valid only as idealisations, or as norms. Economics is based on a formalised variant of common-place folk psychology with individual-level explanation of free agents. It is essential to the research program of economics proper that processes pertaining to whole societies or economies are to be explained by, or reduced to, the behaviour of egocentric, maximising agents.⁹ Usually, it is assumed that comparability of preferences or aggregability of utilities across different individuals does not make sense. It is important, however, that comparisons and aggregation are taken to be possible across different "aspects" or "criteria" of goods, as well as across different potential situations in which the same goods would be available (to various extents). This is part of ordinal and expected utility theory, and I suspect that this may ultimately be the reason for the economists' insistence on a person's preferences being representable by a utility function.

In order to find out whether there is anything economic (referring to economics in this abstract standard sense) about theory revision, we will have to look for patterns of rational choice, i.e., choice that is maximising with respect to some underlying preferences, or more exactly, to some weak ordering which can in turn be represented by a utility function.

3. A second view of economical behaviour: "Informational economy"

There is an alternative and perhaps more intuitive concept of "acting economically" that has played an important role in the development of the research on theory revision. Theory revision is often thought to be *economical* behaviour rather than *economic* behaviour. In the English language, there is a division of labour between the adjectives "economic" and "economical". The former is closely tied to the noun "economics" and means either "of or referring to economics" or has "profitable", "remunerative" and "gainful" as potential synonyms. The adjective "economical", on the other hand, is closely tied to the noun "economy" and means essentially the same as "thrifty", "frugal" or "not wasteful". This particular difference of meanings turns out to be useful for our discussion.

The perspective of *economy* (rather than *economics*) was forcefully taken in Peter Gärdenfors's enormously influential book *Knowledge in Flux* (1988). A glance at the index of the book makes immediately clear that the *criterion of informational economy* is supposed to motivate the essential

⁹ "Egocentric" is not meant to imply "selfish" here. Egocentric agents only look at their personal preference, but the preferences themselves may embody all kinds of (possibly altruistic) thoughts and feelings.

part of the formal modellings of Gärdenfors and his collaborators Carlos Alchourrón and David Makinson. Gärdenfors employs this criterion for the motivation of belief expansions (p. 49), belief revisions (pp. 53, 58) and belief contractions (p. 61).¹⁰ Basically, the criterion is taken to be identical with the idea of *minimal change* (p. 53) and the *conservativity principle* (p. 67). This is very close to Schlick's interpretation of Neurath. According to Gärdenfors,

The key idea is that, when we change our beliefs, we want to retain as much as possible of our old beliefs – information is in general not gratuitous, and unnecessary losses of information are therefore to be avoided. (Gärdenfors 1988, p. 49, similarly on pp. 16, 157)

Ever since the appearance of Gärdenfors's book, the criterion of informational economy has been taken to be a "hallmark" of the paradigm developed by Alchourrón, Gärdenfors and Makinson.¹¹ There is, however, reason to ask why exactly a rational person should be conservative. Gärdenfors's argument that information is not gratuitous does not seem to be sufficient, because information, even if costly, can be wrong, and even if it is correct, it can be misleading. Spelt out a little more explicitly, the argument for conservatism seems to be this:

You shouldn't give away what is valuable

What you have is valuable

You shouldn't give away what you have

The first premise is hardly controversial, it might even be called analytically true as a prescription of economic rationality.¹² The second premise, on the other hand, is much harder to substantiate. It (or at least the idea that it is true) may have evolutionary reasons – probably we would not be alive any more if too many of the statements we accept as true were wrong. Therefore it is at least likely that many of the statements that we happen to hold true now have some survival value.¹³

But the argument for conservatism is a far cry from waterproof. Problems are not hard to come by. Isn't it all too obvious that what we have is not always best? So why should we care to preserve it? It is true that we lose *information* or *content* when we give up some statements of our theory, but it is not clear whether we lose some *truths*, and no-one should object to losing falsities. Notice that

¹⁰ Throughout this paper, it is presupposed that *revisions* have to be *successful* in the sense that they efficiently incorporate the specified new piece of information into the current theory. *Contractions* are called *successful* if they efficiently remove some specified statement from the theory (unless that statement happens to be a logical truth).

¹¹ See, for instance, Boutilier (1996, pp. 264-265) and Darwiche and Pearl (1997, p. 2).

¹² The first premise neglects that one may be forced to give away what is valuable, or that one may invest something at a given time in order to make profit later on.

¹³ I grant that this train of thought is not very compelling. Alternatively, the second premise might be replaced by another one which, however, is at least as dubious: "It is always better to have something than to have nothing."

there is a basic tension here between the economical and the economic precept: The former tends to recommend leaving everything as it is, while the latter recommends striving for the best.

4. Economic and economical considerations in the theory of theory change

In asking what is economic about theory change, we have to keep in mind two different aspects. After the choice-preference-utility line of thinking that we sketched in Section 2, we have found a second one in the thrifty clinging to the sentences one has accepted.

In the work of AGM, it is comparatively easy to recognise the criterion of informational economy at work in expansions of theories by statements that do not contradict the prior theory. AGM recommended to add the new statement and take the deductive closure.¹⁴ However, as we shall see, there are no traces of the criterion for the theory-contravening case which, after all, is the case for which logical models of theory change have primarily been devised. But Gärdenfors's argument can be generalised. In my view, one of the most important philosophical advances of belief revision theory in the last decade is the clear recognition that *belief states* cannot be represented properly by *theories* only. Something else has to be added, namely, some mechanism that tells the agent how to revise her theory. Typically this mechanism exploits some sort of selection function or preference relation.¹⁵ Let us use a neutral name and call the structure exploited by the mechanism a *belief-revision guiding structure*.¹⁶ This argument can be translated into the language of theory change. It then says that scientists do not only subscribe to the theories that they accept, but also to dispositions how to change them should they turn out to be wrong (i.e., roughly, scientists subscribe to a Lakatosian positive heuristics). It turns out that if the new information is inconsistent with the presently accepted theory, it is impossible to apply the pure idea of informational economy on the level of statements. But the idea can be applied on the level of revision-guiding structures. This at the same time defines a form of conservatism that is usable for iterated theory change. Fig. 1 gives a representation of the many senses that "economic" and "economical" can take in theory revision contexts.

¹⁴ In the following, a "theory" is meant to be a set of statements that includes its own logical implications. Thus we endorse the logician's idealisation or stipulation that theories be deductively closed.

¹⁵ Sometimes use is made of use a partitioning or lumping together of the informational contents of the beliefs .

¹⁶ Belief-revision guiding structures typically encode *more* information than the set of currently accepted sentences. They in general allow to retrieve the theory, so that a second component specifying the belief set is unnecessary. Therefore, it is possible to formally identify the result of scientific theorising with a theory-revision guiding structure.

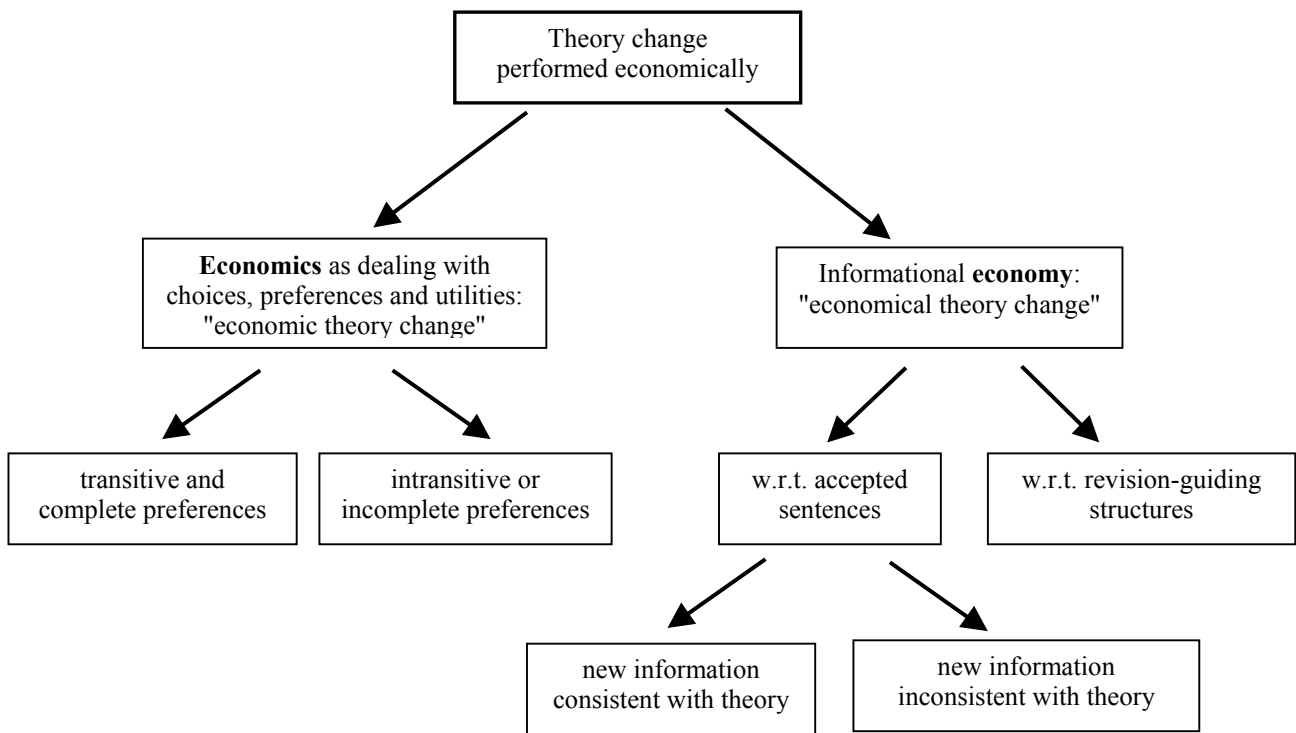


Figure 1

The branch at the left indicates the idea that economic agents are rational or coherent choosers. Though this has not been the principal motivation of theory revision models, we will later see that exactly this idea can be located in the realm of theory formation. On the right-hand branch, we have the idea of informational economy that has always been advertised as the prime driving force of the theory of theory revision (even if thrift in itself does not play a major role in the science of economics). In the next sections, we will trace a few important distinctions within the principle of informational economy, also known as principle of minimal change or conservatism. The maxim of informational economy with respect to sentences tells us: "Don't give away accepted sentences beyond necessity!" The maxim of informational economy with respect to revision-guiding structures tells us: "Don't change your preferences beyond necessity!" These two maxims can further be distinguished as to whether they apply to the theory-contravening case (the one for which theory revision models have originally been invented) or only to the case where the new information is consistent with the agent's theory.

Before entering the systematic discussion, let us summarise some history. A basic form of theory change represents methods that recognise neither economical constraints of minimal change nor economic constraints concerning the rationality of choices. It is possible to add elements of conservatism and elements of rational choice independently from one another. In their seminal work in the 1980s, AGM provided for pure conservatism with respect to sentences (so-called *maxichoice* contraction and revision functions) and for completely unconstrained choices (the case

where only their six *basic* rationality postulates hold).¹⁷ However, they rejected the idea of maxichoice functions as intuitively inadequate, and I think it is fair to say that the main elegance and force of their theories derives precisely from the *supplementary* postulates that go beyond their basic case. So what distinguishes the work of AGM is, on the one hand, a rather strong concept of rational choice generated by transitive and connected preferences, and on the other hand a rather weak concept of conservatism. They provided for conservatism with respect to accepted sentences in the consistent case (where revision reduces to set-theoretic addition plus logical closure), but they provided neither for conservatism in the theory-contravening case nor for conservatism with respect to revision-guiding structures (they provide no change mechanisms for revision-guiding structures at all). So contrary to first appearances, AGM paid a lot more respect to ideas found in economics than to the idea of informational economy.

Systematic variations of the strength of the relevant ideas were investigated only in the 1990s. On the one hand, *weakenings* of AGM's strong presuppositions with respect to the rationalisability of choices are suggested by Lindström (1991) and Rott (1993, 2001).¹⁸ On the other hand, *strengthenings* of the very weak concept of conservatism in AGM are investigated by Boutilier (1993, 1996) and Rott (2003a). Interestingly, it has turned out that the most efficient way – and perhaps the only way – to implement conservatism with respect to accepted sentences in the theory-contravening case is at the same time a form of conservatism with respect to revision-guiding structures. This variant of conservative theory change has actually first been discussed as a particular strategy of extending the classical AGM model so as to equip it with means for performing iterated theory changes. While Boutilier did that in a context that presupposes the full strength of AGM's requirements for rational choice, Rott lifts this restriction and shows that the conservative method can be brought to bear without any assumptions about the coherence of the choices involved.

This little sketch of some developments in the theory change literature shows that the concept of *economical* theory change is independent of the concept of *economic* theory change. In the sense specified, AGM theory revision (minus their two axioms for revisions by inputs that are consistent with the prior theory¹⁹) is not at all economical but, thanks to their "supplementary" postulates, fully economic; the conservative approach investigated in Rott (2003a) is very economical, but not at all economic.

¹⁷ See Alchourrón, Gärdenfors and Makinson (1985).

¹⁸ For similar projects in the related field of non-monotonic reasoning, see Schlechta (1996) and Lehmann (2001).

In the following sections, we recapitulate relevant parts of the actual development of the theory of theory revision, and then discuss the merits and problems of changing theories economically. I will first argue that conservatism with respect to *accepted sentences* has hardly been followed in classical AGM-style theory revision models, and that it is doubtful indeed whether it would be a good idea to follow it. Conservatism with respect to theory-revision guiding structures has been studied as *one* strategy among others in iterable theory change, but it has turned out that it should not be followed either. On the economics side, fairly severe constraints on rational choices have been endorsed in (iterable extensions of) classical AGM-style theory revision. It is true that the AGM postulates can be liberalised systematically according to one's wishes. But we will show that even more modest postulates are problematic – just as problematic as even the basic constraints on rational choices are. We shall see that preferences appear to be context-dependent, giving rise in particular to the formidable problem frequently labelled "the informational value of the menu".

5. Informational economy with respect to theories

In discussing informational economy with respect to theories, we will keep on making two important idealisations: We stipulate that the set of statements accepted by a scientist be logically consistent and closed. This requirement is, of course, wrong as a description of the set of sentences that scientists would affirm to when queried in an interview. However, if we consider a theory to be the set of sentences that we *ascribe* to a scientist from a third person perspective, or the set of sentences that the scientist is *committed to*, then the ideas of consistency and closure lose their implausible appearance. Another simplification we want to make is that when a theory is revised in response to some new piece of information, the revision process in fact successfully incorporates the new information, so that it is accepted in the revised theory. We were taught by Neurath that this is not always reasonable in realistic theory change situations, but I think we can disregard the complications for our discussion. So let us call a consistent and logically closed theory that includes a new piece of information α a *candidate revision* of theory T by α . A (candidate) revision is called *theory-contravening*, if α is inconsistent with the prior theory T .

Let us discuss two attempts at capturing the idea of informational economy on the level of theories:

- (1) When accepting a new piece of information, a scientist should aim at a minimal change of his old theory.

¹⁹ In the common numbering of axioms, these are the third and fourth AGM postulates.

- (2) If there are different ways of effecting the theory change, the scientist should give up those statements that are least entrenched.

These two maxims have frequently been called upon as the principal motivation for logical models of theory change. However, in their most straightforward readings, they are a caricature of what has really been done in the development of the standard models of theory revision. I have argued for this in Rott (2000), and briefly recall the results presented in that paper.

As regards maxim (1). It is possible to show that no two distinct theory-contravening candidate revisions of a consistent and logically closed theory by a sentence α can be set-theoretically compared in terms of the sets of sentences on which they differ with the prior theory.

For the discussion of maxim (2), we need a little bit of terminological preparation. A statement α is *more entrenched* in a theory than another statement β if and only if the agent holds on to α and gives up β when learning that not both α and β are (or may be) true.²⁰ A new piece of information α is called *moderately surprising* if $\neg\alpha$ is a non-minimal element of the prior theory with respect to epistemic entrenchment. Let us call a revision by α *amnesic* if the revised theory consists of nothing else but $Cn(\alpha)$ (where Cn is some ordinary well-behaved Tarskian logic); otherwise we call the revision *anamnesic*. Now suppose we want to revise a theory by a sentence α and identify two elements of the initial theory that non-redundantly entail $\neg\alpha$. Then it may well happen, when performing a regular AGM-style revision by α , that the scientist removes the *more* entrenched and retains the *less* entrenched statement. An instance of such a situation can indeed *always* be identified when α is moderately surprising and the revision by α is anamnesic. Thus neither of the two maxims that have always belonged to the core rhetoric accompanying AGM-style theory change models is actually obeyed in these very models.

This observation may seem too baffling at first sight, but it has turned out to be rather robust. I have argued in Rott (2000) that it does not help to talk about theory contraction in place of theory revision, that it does not help to have recourse to AGM's celebrated representation results that link theory change to some sort of minimisation process, and that it does not help either to restrict attention only to those statements of a theory that are actually true. Yet another route to saving the idea of minimal change can be taken by holding that informational economy should not pertain to theories as sets of statements, but to richer representations of the results of scientific theorising: to dispositions for theory change, or equivalently, to structures that are suitable to guide (iterated)

²⁰ See Gärdenfors and Makinson (1988) and Rott (2001).

theory revisions. We will presently address this point in some detail. Before doing this, however, we have a look at potential norms for theory change.

First, we have to be aware of the fact that the demand for informational economy conflicts with other desiderata. For instance, it competes with the *synchronic* or *static coherence* constraints of *logical consistency* and *closure*. We should give up something when we find ourselves caught in an inconsistency, and we typically have to give up not only the culprit sentence, but also many of the sentences that could be derived with the help of the culprit. This is, among other things, because we want to maintain the logical closure of our theories. Of course, a change that is minimal subject to the constraints of consistency and closure will in general be bigger than a change that is minimal when no constraints are to be respected. Conservativity is itself a criterion of *diachronic* or *dynamic coherence* (Rott 1999). There are more concepts of coherence that we will have reason to consider below, viz., *dispositional coherence* and *temporal coherence*. At this juncture it is not at all clear whether the latter two concepts might generate requirements that compete with the requirements of informational economy.

Secondly, it is instructive to contrast the ideas that are advocated in the logical literature on theory revision with ideas recommended by philosophers of science who think about theory change as arising in the (r)evolution of scientific theories and research programmes. In their elementary but thought-provoking little book *The Web of Belief*, Quine and Ullian (1978) list six 'virtues' of hypotheses: empirical adequacy, simplicity, conservatism, modesty, generality, refutability and precision. Logical models of theory change, on the other hand, have appealed to success, logical closure and consistency, informational economy, dispositional coherence, preference and indifference,²¹ temporal coherence (these terms will be explained in due course.) A simple check of these lists reveals that the only criterion that is endorsed both by Quine and Ullian and the standard account of theory revision is informational economy. And a moment's thought makes it clear that at least some of the virtues mentioned compete with one another. Simplicity often requires deviations from informational economy, modesty conflicts with refutability.

Fortunately, we do not have to commit ourselves to one or the other set of criteria here. It is enough at this juncture to note that an account of how to integrate various criteria of *theory choice* (the perspective dominant in the philosophy of science) and *theory change* (the main perspective studied in philosophical logic) is badly needed, but has never been offered. In any case, there is no obvious reason for according informational economy a privileged status among the many contenders that we have identified. It has only a very restricted normative force.

²¹ Pagnucco and Rott (1999).

6. Conservatism with respect to theory-revision guiding structures

We now turn to another interpretation of the idea that scientists should aim at preserving what they have. The propositional content encoded in statements of the scientific object language is not the only kind of valuable information. We may also be interested in preserving non-propositional information encoded in theory-revision guiding structures, i.e., in richer representations of the results of scientific theorising.

Rich representations of theoretical structures that have gained some currency in the literature include *plausibility orderings* of possible worlds. Given such an ordering, the core set of the most plausible worlds contains those worlds that are consistent with the current scientific theory, i.e., those that could be the real world given what the scientist knows. The ordering of the remaining worlds reflects their relative distance from this core set. Theory change prompted by some new piece of information α then proceeds by manipulating the ordering of worlds in such a way that all the most plausible worlds in the revised ordering satisfy α . This constraint is mild and leaves a lot of freedom for the exact specification of a coherent revision mechanisms. The most *conservative* or *economical* way of changing the plausibility ordering that respects this constraint was defined and investigated by Boutilier (1993, 1996).

Another way of richly representing theoretical structures consists in *entrenchment orderings* of sentences (their interpretation was already mentioned in the previous section). Given an entrenchment ordering, the statements that are least plausible under this ordering are those sentences that are not accepted at all. The ordering of the remaining statements reflects how firmly they are endorsed by the scientist, i.e., the relative tenacity with which the scientist is determined to cling to his statements. In this model, theory change prompted by a new piece of information α proceeds by manipulating the ordering of sentences in such a way that α is not among the least entrenched sentences under the revised ordering. Again, this constraint is a rather mild one and leaves a lot of room for filling in the details on an exact specification of the revision mechanism. The most *conservative* or *economical* way of changing the entrenchment ordering that respects this constraint is investigated in Rott (2003a).

It turns out that the account based on entrenchment orderings is essentially a generalisation of the possible worlds account proposed by Boutilier. In the specific context of the assumptions made by Boutilier, however, the methods are equivalent. Both methods satisfy an axiom for iterated theory change which is sufficient to characterise conservative revisions of richer representations of

theories or theoretical structures. As shown by Boutilier (1993, 1996) and, in a more general setting, by Rott (2003a), a repeated conservative change of T first by a statement α and then by a second statement β leads to the same result as a single conservative revision of T by β , provided that β is inconsistent with the result of the revision of T by α . In symbols: $(T*\alpha)*\beta = T*\beta$ if β is inconsistent with $T*\alpha$. Why do we call this recipe "conservative"? Because it suggests that if β cannot be accommodated consistently, the way of handling it in the revised theory $T*\alpha$ is just the same as it was in the original T – as if the agent had never learnt about α . Loosely speaking, the structure of the old theory is stronger than the new piece of evidence α .

Almost immediately after Boutilier had suggested conservative theory change as a natural extension of the AGM model, Goldszmidt and Pearl (1994, 1997) discovered that the behaviour defined by this model is intuitively queer. They gave the following example. A person who we may for the sake of argument take to be in a state of complete ignorance observes an animal that she takes to be a bird (b). As the animal comes closer, the person perceives that the animal is red (r). A few moments later, she realises (perhaps informed by an expert) that the animal is not a bird after all ($\neg b$). If we use Boutilier's method of conservative theory change, the result of these three subsequent revisions is $\text{Cn}(\neg b)$.

Goldszmidt and Pearl rightly argue that this is not satisfactory. Why "forget" about the colour of the animal just because it turns out that the person has misclassified the animal as a bird? In more general terms, it can be shown that this model is *temporally incoherent*. The AGM-postulate of "success" for revisions says – *pace* Neurath – that the most recent piece of information should always be included in the revised theory. Thus, at the moment of receipt the most recent piece of information is being treated as the "most important" one. But this privilege is immediately lost when new information happens to come in. To see this, let α , β and γ stand for sentences that are pairwise consistent, but jointly inconsistent. Then iterated conservative change of the trivial theory $\text{Cn}(\emptyset)$ first by α , then by β and finally by γ results in the theory $\text{Cn}(\alpha, \gamma)$.²² The first and the last piece of information are stronger than the one that comes in between.

Conservatism with respect to revision-guiding structures thus has unacceptable consequences when applied as a method for iterated theory revision. It violates the requirement that a good method of

²² Both AGM and Boutilier's models assume that the orderings of worlds or sentences involved are complete preorderings, i.e., that all worlds and all sentences are comparable as regards their plausibility or entrenchment. So it may perhaps be suspected that the problem of the conservative revision strategy is (at least in part) due to the strong requirements of "dispositional coherence" inherent in the AGM model upon which Boutilier's model is built. However,

theory change be *temporally coherent*, i.e., coherent in its attitude towards the value of the recency of information. As long as one decides invariably to accept new information (i.e., to take the last piece of information to be the most important one), the only coherent attitude towards the recency of information is to take the second-last piece of information to be the second-most important one, and so on. Instead of $Cn(\alpha, \gamma)$ as above, the desired result would thus be $Cn(\beta, \gamma)$. There is in fact an alternative model that yields precisely this result (first studied systematically by Nayak 1994).²³ But as we said before, Neurath's principle that leaves the scientist the free choice whether or not to accept a new sentence (as a protocol statement) is better from a philosophical point of view. Being altogether indifferent towards the time at which a certain piece of information is received is a form of temporal coherence, too.

We now leave the field of "economic reasoning" and turn to "economical reasoning", i.e., to the left branch of Figure 1 that refers us to the realm of choices, preferences and utilities.

7. Rational choices and logical properties

As pointed out above, the classical AGM model of theory revision embodies a strong idea of dispositional coherence. More precisely, the agent's dispositions to change his theory in potential revisions are guided by choices that can be rationalised by a complete preordering, i.e., the agent acts *as if* he was a maximiser with respect to this ordering. All possible worlds are presumed to be comparable with each other in terms of plausibility (AGM 1985, Grove 1988), all statements are presumed to be comparable with each other in terms of entrenchment (Gärdenfors and Makinson 1988). These facts find expression in the seventh and eighth postulates of Alchourrón, Gärdenfors and Makinson (1985) which constrain the scientist's disposition to change his theory:

$$(*7) \quad (T * \alpha) \cap (T * \beta) \subseteq T * (\alpha \vee \beta)$$

which is sometimes called "Disjunction in the antecedent", and

$$(*8) \quad \text{If } \neg\beta \notin T * \alpha, \text{ then } T * \alpha \subseteq T * (\alpha \wedge \beta)$$

which is sometimes called "Rational monotonicity". These postulates have turned out to correspond to conditions on semantic or syntactic choice functions known as *Sen's properties* α and β^+ which in effect require that the revision function be rationalisable by a complete and transitive preference

the unwelcome effects of temporal incoherence remain present in exactly the same way even if *all* of the dispositional requirements of the AGM model are dropped (Rott 2003a).

²³ I discuss the merits of this model of "moderate belief revision" in some detail in Rott (2003a). Also compare Nayak, Pagnucco and Peppas (2003).

relation. Many concrete systems of theory revision do not satisfy (*7) and (*8) – a fact that shows that these AGM axioms are very strong. Knowing this, it is pleasing to find that one can draw on the rich resources of the theory of rational choice in order to introduce appropriate weakenings of the theory revision postulates. For instance, a weaker form of (*8) is

(*8c) If $\beta \in T * \alpha$, then $T * \alpha \subseteq T * (\alpha \wedge \beta)$

also called "Cumulative monotonicity". This condition corresponds to "Aizerman's axiom" in the theory of rational choice (see Moulin 1985). The theory of rational choice is a powerful instrument which is suitable for analysing and constructing revision operations that are much more flexible than the original AGM ones. Rott (2001, Chapter 7) describes how exactly one can use semantic choice functions (for the selection of the most plausible worlds) and syntactic choice functions (for the selection of the least entrenched sentences) in the construction of theory revisions, and also how postulates for theory revision correspond to rationality requirements for semantic and syntactic choice functions.²⁴ Seen from this perspective, the theory of theory change can indeed be interpreted as being based on economical principles.

Almost from its beginning, however, the classical theory of rational choice has been subjected to serious criticism. In the rest of this section we present an argument to the effect that a fundamental problem for the theory of rational choice transfers directly to theories of theory revision.

Consider the following example which is not a scientific one in order to keep things simple. A well-known philosophy department has announced a position in metaphysics. Among the applicants for the job there are a few persons we happen to know. First, there is Amanda Anderson, a highly profiled, excellent metaphysician. Second, we have Bernice Becker, who is also definitely very good, though not quite as highly profiled, in metaphysics as Anderson. In addition, Becker has done some substantial work in logic. A third applicant is Carlos Cortez. He has a comparatively slim record in metaphysics, but he is widely recognised as one of the most brilliant logicians of his generation.

Suppose that our initial "theory" about the case – i.e., our initial set of beliefs and expectations – includes that neither Anderson nor Becker nor Cortez will get the job (say, because we think that Derek Davidson, an outstanding metaphysician, is the obvious candidate who is going to get the position anyway). Let us also be clear about the fact that there is only one job available.

²⁴ Important earlier work was done by Lindström (1991).

Consider now three hypothetical scenarios, each of which describes a *potential* development (*not* a *sequence* of stages) of the selection procedure. In each of these alternative scenarios we are genuinely taken by surprise, because we learn that one of the candidates we had believed to be losing will be offered the position. To make things shorter, we introduce some abbreviations. Let the letters a , b and c stand for the statements that Anderson, Becker and Cortez, respectively, will be offered the position.

Scenario 1. The dean tells us in confidence that it has been decided that either Anderson or Becker will be appointed.²⁵ This message comes down to supplying us with the premise $a \vee b$. Given this premise, we conclude that Anderson, being the better metaphysician, will get the job. We also infer that the other candidates will return empty-handed.

Scenario 2. This is a very unexpected scenario in which we are told by the dean that Cortez is actually the only serious candidate left in the competition. Fortunately, there is no need to invest a lot of thinking here. We accept c in this case.

Scenario 3. In this scenario the dean tells us that it has been decided that either Anderson or Becker or Cortez will get the job, thus supplying us with the premise $a \vee b \vee c$. This piece of information triggers off a rather subtle line of reasoning. Knowing that Cortez is a splendid logician, but that he can hardly be regarded as a metaphysician, we realise that competence in logic is considered to be a non-negligible asset by the selection committee. Still we keep on believing that Cortez will not make it, because his credentials in metaphysics are just too weak. Since, however, logic appears to contribute positively to a candidate's profile, we conclude that Becker, and not Anderson, will get the job.

This qualitative description should do for our purposes, but for readers who prefer more precision, the following story may help. The selection committee has decided to assign points to evaluate the candidates' work. Anderson scores 97 out of 100 in metaphysics, but has done no logic whatsoever, so she scores 0 here. Becker scores 92 in metaphysics and a respectable 50 in logic. Cortez scores only 40 in metaphysics, but boasts of 99 in logic. In scenario 1, we take it that metaphysics is the only criterion, so clearly Anderson must be the winner. Scenario 2 is trivial. In scenario 3, we gather that, rather unexpectedly, logic matters. As can easily be verified, any weight we attach to

²⁵ We take it for granted in this example that the dean is not playing games with us, that she is not lying, that she has the relevant knowledge etc.

logic between $1/10$ and $1/2$ (with metaphysics taking the rest) will see Becker end up in front of both Anderson and Cortez.

Let us now summarise our conclusions from the various premises that the dean supplies us with. Our initial theory T contains $\neg a$, $\neg b$, $\neg c$ and d among other things. In scenario 1, the new piece of information $a \vee b$ leads us to accept a and $\neg b$ (along with $\neg c$ as well as $\neg d$ which we will not mention any more). In scenario 2, c simply makes us accept that $\neg a$ and $\neg b$. In scenario 3, the new piece of information $a \vee b \vee c$ leads us to accept that $\neg a$ and b . We can now show that this situation refutes some of the basic logical principles of "economic" theory revision.

First, the example shows that "disjunction in the premises" (*7) does not hold. Take (*7) and substitute $a \vee b$ for α and c for β . Then notice that $\neg b$ is believed if the input is $a \vee b$, and also if the input is c . But $\neg b$ is not believed if the input is $a \vee b \vee c$. Thus the revised theory $T^*(a \vee b \vee c)$ does not include what is common to $T^*(a \vee b)$ and $T^* c$, and (*7) is violated.

Secondly, we find that the situation does not conform to the weakened monotonicity postulate (*8c). Take (*8c) and substitute $a \vee b \vee c$ for α and $a \vee b$ for β . Even though we believe that $a \vee b$ is true if we are given the information $a \vee b \vee c$, it is not the case that everything believed on the basis of the latter is also believed on the basis of $(a \vee b \vee c) \wedge (a \vee b)$ which is equivalent with $a \vee b$. Sentences $\neg a$ and b are counterexamples. Thus the revised theory $T^*(a \vee b \vee c)$ is not a subset of the theory $T^*((a \vee b \vee c) \wedge (a \vee b)) = T^*(a \vee b)$, and (*8c) is violated. *A fortiori*, (*8) is violated as well.

What do these problems derive from? We said that principles of theory change can be systematically interpreted in terms of rational choice. On this interpretation, "disjunction in the premises" (*7) turns out to be an instantiation of one of the most fundamental conditions – perhaps *the* most fundamental condition – of the theory of rational choice: Sen's Property α . This condition, also called *Independence of Irrelevant Alternatives* or *Chernoff property*, says that any element which is optimal in a certain menu remains an optimal element after some other elements have been cancelled from the menu. The three scenarios in our example are modelled after well-known choice situations in which Property α is violated – cases that also happen to disobey Aizerman's axiom. Both properties may fail to be satisfied if the very 'menu' from which the agent is invited to choose carries important information. This phenomenon which Sen calls the 'epistemic value' or the

'epistemic relevance of the menu'²⁶ suggests that the context of choice has a decisive influence upon the shape of the preference relation of the agent – an idea strongly opposed to the idea of context-independent preferences that underlies the classical theory of rational choice. The *locus classicus* for the problem is a passage in Luce and Raiffa (1957, p. 288) who chose to avoid the problem of the epistemic value of the menu by *fiat*:

This illustrates the important assumption implicit in axiom 6 [essentially Property α , H.R.], namely, that adding new acts to a decision problem under uncertainty *does not alter one's a priori information as to which is the true state of nature*. In what follows, we shall suppose that this proviso is satisfied. In practice this means that, if a problem is first formulated so that the availability of certain acts influences the plausibility of certain states of nature, then it must be reformulated by redefining the states of nature so that the interaction is eliminated.

This may make good sense as a rejoinder in the context of the general theory of choice and decision. An explanation of how information is surreptitiously conveyed through the particular contents of the menu and how it affects the chooser's preferences is simply not this theory's business. Unfortunately, the same defence is not available for the problem highlighted by the above job example. It *is* the task of the theory of theory formation to model how one's prior theory is affected by information received from external sources. This is precisely what this theory has been devised to explain, and therefore the anomaly cannot be pushed to a neighbouring research field.

The question raised by our example is a general one. It is hard to get rid of the feeling that the dean's information about the final candidates conveys more information than meets the ear. The fact that a logician gets *mentioned* as a top-ranking contender or that logic becomes a *topic* seems to carry surplus information, over and above the propositional content of the corresponding statement. Does the very fact that a statement is offered in a menu for acceptance have a special relevance for processes of theory revision that has been overlooked so far? Or are there other ways out of the predicament? In any case, the upshot of this section is that the economist's point of view supplies some solid planks for Neurath's boat, but by the same token, it also makes us suspect that these planks cannot be put together without the boat springing a leak.

8. Conclusion

This paper started out from Neurath's coherence theory of knowledge and Schlick's claim that Neurath makes implicit use of an economy principle to alleviate the notorious ambiguities inherent in coherentist epistemology. We then reviewed work in the tradition of the AGM approach to

²⁶ Sen (1993, pp. 500–503; 1995, pp. 24–26) has brought the problem to wide attention.

theory revision, arguably the most prominent logical paradigm of purifying theories from contradictions. Our overarching questions were to what extent economic(al) principles have played a role in the *actual* development of this paradigm, and to what extent such considerations *should* have been followed. Our conclusions are mostly negative. Informational economy (conservatism with respect to accepted sentences), although widely advertised as *the* central motivation of theory revision models, turns out not to have played anything like a dominant role in the development of such models, and we have found no reason why it should. I think that Schlick was right about that. Conservatism with respect to revision-guiding preferences has fact been suggested as a strategy for iterated theory revision, but it soon turned out to have unwelcome consequences. So the theory of theory revision has as a matter of fact not focused on *economy*, and the idea of *economical* theory changes has very limited normative force, too. As regards *economic* theory change, our findings are more encouraging. It is possible to reconstruct large parts of theory revision in terms of rational choice theory. So as a matter of fact, ideas coming from *economics* have prevailed in the AGM paradigm and related approaches. However, at the end of the paper we found that a fundamental problem of the general theory of choice seriously infects the specific application area of theory change. Thus the use of rational choice theory has helped us to spot a new puzzle rather than solve old problems. There is a lot of work in cognitive economics that waits to be done.²⁷

References

- Alchourrón, Carlos, Peter Gärdenfors and David Makinson: 1985, 'On the Logic of Theory Change: Partial Meet Contraction Functions and Their Associated Revision Functions', *Journal of Symbolic Logic* 50, 510-530.
- Boutilier, Craig: 1993, 'Revision Sequences and Nested Conditionals', in R. Bajcsy (ed.), *IJCAI-93 - Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, 519-525.
- Boutilier, Craig: 1996, 'Iterated Revision and Minimal Change of Conditional Beliefs', *Journal of Philosophical Logic* 25, 263-305.
- Broome, John: 1999, *Ethics Out of Economics*, Cambridge University Press, Cambridge.
- Darwiche, Adnan, and Judea Pearl: 1994, 'On the Logic of Iterated Belief Revision', in Ronald Fagin, ed., *TARK '94 - Proceedings of the Fifth Conference on Theoretical Aspects of Reasoning About Knowledge*, Pacific Grove, Cal.: Morgan Kaufmann, 5-23.
- Darwiche, Adnan, and Judea Pearl: 1997, 'On the Logic of Iterated Belief Revision', *Artificial Intelligence* 89, 1-29.
- Debreu, Gerard: 1959, *Theory of Value*, Wiley, New York.
- Edgeworth, Francis Ysidro: 1881, *Mathematical Psychics*, Kegan Paul, London.
- Gärdenfors, Peter: 1979, 'Conditionals and Changes of Belief', in Illkka Niiniluoto and Raimo Tuomela (eds.), 'The Logic and Epistemology of Scientific Change', *Acta Philosophica Fennica* 30 (1978), nos. 2-4, 381-404.

²⁷ Most of the material presented here is discussed in more technical detail in Rott (2000, 2001, 2003a, 2003b). Thanks go to Vincent F. Hendricks, Eckehart Köhler, Hannes Leitgeb, Nils-Eric Sahlin, Gerhard Schurz, Krister Segerberg and Wolfgang Spohn for very helpful comments and discussions.

- Gärdenfors, Peter: 1988, *Knowledge in Flux. Modeling the Dynamics of Epistemic States*, Bradford Books, MIT Press, Cambridge, Mass.
- Gärdenfors, Peter, and David Makinson: 1988, 'Revisions of Knowledge Systems Using Epistemic Entrenchment', in Moshe Vardi (ed.), *TARK'88 - Proceedings of the Second Conference on Theoretical Aspects of Reasoning About Knowledge*, Los Altos: Morgan Kaufmann, 83-95.
- Grove, Adam: 1988, 'Two Modellings for Theory Change', *Journal of Philosophical Logic* 17, 157-170.
- Haller, Rudolf: 1982a, 'New Light on the Vienna Circle', *The Monist* 65, 25-35.
- Haller, Rudolf: 1982b, 'Das Neurath-Prinzip – Grundlagen und Folgerungen', in Friedrich Stadler (ed.), *Arbeiterbildung in der Zwischenkriegszeit*, Löcker, Wien, 79-87.
- Hansson, Sven Ove: 1999, *A Textbook of Belief Dynamics: Theory Change and Database Updating*, Kluwer, Dordrecht.
- Hansson, Sven Ove (ed.): 1997, Special Issue on 'Non-Prioritized Belief Revision', *Theoria* 63, 1-134.
- Hausman, Daniel: 1998, 'Economics, Philosophy of', *Routledge Encyclopedia of Philosophy*, electronic edition.
- Hausman, Daniel M., and Michael S. McPherson: 1996, *Economic Analysis and Moral Philosophy*, Cambridge University Press, Cambridge.
- Lehmann, Daniel: 2001, 'Nonmonotonic Logics and Semantics', *Journal of Logic and Computation* 11, 229-256.
- Lindström, Sten: 1991, 'A Semantic Approach to Nonmonotonic Reasoning: Inference Operations and Choice', Uppsala Prints and Preprints in Philosophy, Department of Philosophy, University of Uppsala, 1991:6.
- Luce, R. Duncan, and Howard Raiffa: 1957, *Games and Decisions*, Wiley, New York.
- McFadden, Daniel: 1999, 'Rationality for Economists?', *Journal of Risk and Uncertainty* 19, 73-105.
- Moulin, Hervé: 1985, 'Choice Functions over a Finite Set: A Summary', *Social Choice and Welfare* 2, 147-160.
- Nayak, Abhaya: 1994, 'Iterated Belief Change Based on Epistemic Entrenchment', *Erkenntnis* 41, 353-390.
- Nayak, Abhaya, Maurice Pagnucco and Pavlos Peppas: 2003, 'Dynamic Belief Revision Operations', *Artificial Intelligence*, in press.
- Neurath, Otto: 1931, 'Soziologie im Physikalismus', *Erkenntnis* 2, 393-431; English translation 'Sociology in the Framework of Physicalism' in O.N., *Philosophical Papers 1913-1946*, eds. Robert S. Cohen and Marie Neurath, Reidel, Dordrecht 1983, pp 58-90.
- Neurath, Otto: 1932/33, 'Protokollsätze', *Erkenntnis* 3, 204-214; English translation 'Protocol Statements' in O.N., *Philosophical Papers 1913-1946*, eds. Robert S. Cohen and Marie Neurath, Reidel, Dordrecht 1983, pp 91-99.
- Neurath, Otto: 1934, 'Radikaler Physikalismus und "Wirkliche Welt"', *Erkenntnis* 4, 346-362; English translation 'Radical Physicalism and the "Real World"' in O.N., *Philosophical Papers 1913-1946*, eds. Robert S. Cohen and Marie Neurath, Reidel, Dordrecht 1983, pp 100-114.
- Pagnucco, Maurice, and Hans Rott: 1999, 'Severe Withdrawal (and Recovery)', *Journal of Philosophical Logic* 28, 501-547. (Corrected reprint in the *JPL* issue of February 2000.)
- Quine, Willard V.O., and Joseph S. Ullian: 1978, *The Web of Belief*, second edition, Random House, New York.
- Rosenberg, Alexander: 1983, 'If Economics Isn't Science, What Is It?', *Philosophical Forum* 14, 296-314. Reprinted in E.D. Klemke et al. (eds.), *Introductory Readings in the Philosophy of Science*, 3rd edition, Prometheus Books, Amherst 1998, 154-170.
- Rosenberg, Alexander: 1995, 'Philosophy of Economics', in Robert Audi (ed.), *Cambridge Dictionary of Philosophy*, Cambridge University Press, Cambridge, pp. 582-583.
- Rott, Hans: 1993, 'Belief Contraction in the Context of the General Theory of Rational Choice', *Journal of Symbolic Logic* 58, 1426-1450.

- Rott, Hans: 1999, 'Coherence and Conservatism in the Dynamics of Belief. Part I: Finding the Right Framework', *Erkenntnis* 50, 387-412.
- Rott, Hans: 2000, 'Two Dogmas of Belief Revision', *Journal of Philosophy* 97, 503-522.
- Rott, Hans: 2001, *Change, Choice and Inference*, Oxford University Press.
- Rott, Hans: 2003a, 'Coherence and Conservatism in the Dynamics of Belief. Part II: Iterated Belief Change Without Dispositional Coherence', *Journal of Logic and Computation* 13, 111-145.
- Rott, Hans: 2003b, 'A Counterexample to Six Fundamental Principles of Belief Formation', manuscript, January 2003.
- Schlechta, Karl: 1996, 'Some Completeness Results for Stopped and Ranked Classical Preferential Models', *Journal of Logic and Computation* 6, 599-622.
- Schlick, Moritz: 1934: 'Über das Fundament der Erkenntnis' in *Erkenntnis* 4, 79-99; English translation 'The Foundation of Knowledge', in Alfred J. Ayer (ed.), *Logical Positivism*, New York, Free Press 1959, 209-227.
- Sen, Amartya K.: 1993, 'Internal Consistency of Choice', *Econometrica* 61, 495-521.
- Sen, Amartya K.: 1995, 'Is the Idea of Purely Internal Consistency of Choice Bizarre?', in J. E. J. Altham and R. Harrison (eds.), *World, Mind, and Ethics. Essays on the Ethical Philosophy of Bernard Williams*, Cambridge University Press, Cambridge, 19-31.
- Spohn, Wolfgang: 2003, 'Die Logik und das Induktionsproblem', in Peter Schroeder-Heister und Wolfgang Spohn (eds.), *Logik in der Philosophie*, Synchron-Verlag, Heidelberg.