

Coherence, Justification, and the AGM Theory of Belief Revision

Michael J. Shaffer

The University of North Carolina at Wilmington

ABSTRACT: In a recent article, Peter Gärdenfors has suggested that the AGM (Alchourrón, Gärdenfors, and Makinson) theory of belief revision can be given an epistemic basis by interpreting the revision postulates of that theory in terms of a version of the coherence theory of justification. To accomplish this goal Gärdenfors suggests that the AGM revision postulates concerning the conservative nature of belief revision can be interpreted in terms of a concept of epistemic entrenchment and that there are good empirical reasons to adopt this view as opposed to some form of foundationalist account of the justification of our beliefs. In this paper I argue that Gärdenfors attempts to underwrite the AGM theory of belief revision by appealing to coherentism is seriously inadequate for several reasons. First, Gärdenfors appeals to empirical studies which he claims show that actual epistemic agents do not keep track of the justification for their beliefs and that such agents are not inclined to cede beliefs based on justifications that are defeated, and, hence, that foundationalist epistemologies are seriously incorrect as descriptions of real epistemic agents. I argue that these experiments show no such thing, and that they reveal little or nothing about how epistemic agents should act. In any case, Gärdenfors argues that we should adopt a coherence theory of justification in light of his criticisms of foundationalist epistemologies. In adopting a coherentist epistemology Gärdenfors argues that we can, in effect, replace the concept of justification with that of epistemic entrenchment. I shall argue that there are serious conceptual problems with this proposal that concern the formal features of justification and entrenchment. Based on these observations, I argue that entrenchment is not a plausible replacement for the concept of justification. Finally, I make some general remarks concerning the AGM theory and coherence theories from the perspective of theories of epistemically rational behavior.

1. Introduction

In a recent article, Peter Gärdenfors (1992) has suggested that the AGM (Alchourrón, Gärdenfors, and Makinson) theory of belief revision can be given an epistemic basis by interpreting the revision postulates of that

theory in terms of a version of the coherence theory of justification. To accomplish this goal Gärdenfors suggests that the AGM revision postulates concerning the conservative nature of belief revision can be interpreted in terms of a concept of epistemic entrenchment and that there are good empirical reasons to adopt this view as opposed to some form of foundationalist account of the justification of our beliefs.

In this paper I argue that Gärdenfors' attempt to underwrite the AGM theory of belief revision by appealing to a form of coherentism is seriously inadequate for several reasons.¹ First, Gärdenfors appeals to empirical studies (collected in Ross and Anderson 1982) which he claims show, or at least suggest, that actual epistemic agents do not keep track of the justification for their beliefs and that such agents are not inclined to cede beliefs based on justifications that are defeated. Hence, he argues that foundationalist epistemologies are seriously incorrect as descriptions of real epistemic agents because they are alleged to require that beliefs be justified exclusively in terms of such evidence; in other words, foundationalists wrongly hold that belief is a matter of pedigree and is to be understood exclusively in virtue of undefeated evidence in the form of beliefs. I argue that these empirical studies show no such thing, and, more importantly, that they reveal little if anything about how epistemic agents should act. In any case, Gärdenfors argues that we should adopt a coherence theory of justification in light of this criticism of foundationalist epistemologies.

Second, and perhaps more importantly, in adopting a coherentist epistemology on the basis of the empirical evidence referred to above, Gärdenfors argues that we can, in effect, replace the concept of justification with a formal concept of epistemic entrenchment. I shall argue that there are several serious problems with this proposal. The first two problems with this suggestion concern the formal features of the concepts of justification and entrenchment, and the basic function of those concepts respectively. More specifically, it is not clear that entrenchment can be coherently understood as admitting of precise and continuous numerically measurable degrees, and it does not seem to be the case that entrenchment is an obviously normative concept in the manner that justification is. Finally, I suggest that Gärdenfors' attempt to replace justification with entrenchment makes no sense in terms of the coherentist's particular concept of justification. Based on these considerations, I argue that entrenchment is not a plausible replacement for the concept of justification, and that Gärdenfors' attempt to underwrite the AGM theory of belief revision by allying it with the coherence theory of justification is a failure. Finally, I make some general remarks concerning the AGM theory and coherence theories from the perspective of theories of epistemically rational behavior.

2. The AGM Theory of Belief Revision: An Overview

Before presentation of these criticisms, it will be instructive to those who are not familiar with the AGM theory of belief revision to take a brief look at this influential theory of belief dynamics.² The purpose of the AGM theory is to provide us with a theory of the rationality of belief change, and this sort of theory stands in sharp contrast to traditional theories of epistemic rationality that almost universally deal with belief support or justification; with the rationality of belief as such. So the AGM theory of belief revision ought to be viewed as a theory of the dynamics of belief states, as opposed to a theory of the static features of belief states. The essential idea behind this epistemological theory is that there are normative rules that govern how one ought to change one's beliefs in light either of acquiring new beliefs or revising beliefs that one already holds.

The AGM theory is fundamentally based on the concept of a belief state, belief set or a corpus of beliefs, K , satisfying the following minimal conditions where it is assumed that belief states are given a representation in some language \mathcal{L} and where a, b, c, \dots are sentences of \mathcal{L} :

- (DfBS) A set of sentences, K , is a belief state if and only if (i) $\neg(K \vdash \perp)$, and (ii) if $K \vdash b$, then $b \in K$.

Here 'vdash' is a standard definition of logical consequence for \mathcal{L} , and so DfBS requires that K is logically consistent and is closed under logical consequence. We can then define the content of a belief state as the set of logical consequences of K , so $\{b : K \vdash b\} =_{df} Cn(K)$. Given this basic form of epistemic representation, the AGM theory is intended to be a normative theory about how a given belief state K_i which satisfies (DfBS) is related to other belief states K_n satisfying (DfBS) relative to: (1) the addition of a new belief b to K_i , or (2) the retraction of b from K_i , where $b \in K_i$. Belief changes of the latter kind are termed *contractions* ($\dot{-}$), but belief changes of the former kind must be further sub-divided into those that require giving up some elements of K_i and those that do not. Additions of beliefs that do not require giving up previously held beliefs are termed *expansions* ($+$), and those that do are termed *revisions* ($*$).³ As such, the dynamics of beliefs will then simply be the epistemically normative rules that govern rational cases of contraction, revision and expansion of belief states.

Expansion of K_i by a sentence b is defined using the notion of the content of a belief state as follows:

$$(Df+) K + b = Cn(K \cup \{b\}).$$

Moreover, revision can be defined in terms of expansion and contraction in accord with the *Levi identity* as follows:

$$(Levi identity) K * b = (K \dot{-} \neg b) + b.$$

Combining the simple concept of belief state expansion defined above with some principled concept of belief state contraction then yields a complete specification of the possible dynamic belief changes relative to belief states. Contraction obviously then turns out to be the central concept of the AGM theory.

2.1 Contraction, Conservation, and Partial Meet

The fundamental insight behind the AGM theory is that belief changes that are contractions should be fundamentally conservative in nature; in other words, in belief changes one ought to make the minimal alterations necessary to incorporate new information and maintain or restore logical consistency. This fundamental assumption in the AGM theory is supposed to be justified in virtue of a principle of informational economy that holds that information is valuable and so we should retain it at all costs unless we are forced to do otherwise.⁴ Gärdenfors (1992) presents the informal version of this intuition, which he calls the principle of conservation (POC), as follows:

(POC) When changing beliefs in response to new evidence, you should continue to believe as many of the old beliefs as possible (381).

After having considered and rejected several interpretations of this basic intuition in Alchourrón, Gärdenfors and Makinson 1985 and Gärdenfors 1988, the POC was given formal explication in terms of the concept of partial meet contraction (PMC). Partial meet contraction is defined as follows:

(PMC) $K \dot{\div} b = \bigcap \gamma(K \perp b)$.

$K \perp b$ is inclusion-maximal set of subsets of K that do not imply b , γ is a selection function such that $\gamma(K \perp b)$ is a non-empty subset of $K \perp b$, unless b is empty. Where b is empty $\gamma(K \perp b)$ is just K . The so-called AGM postulates explicitly tell us what rules govern such changes, and they are as follows:

(P1-Closure) $Cn(K \dot{\div} b) = K \dot{\div} b$,

(P2-Inclusion) $K \dot{\div} b \subseteq K$,

(P3-Vacuity) If $b \notin K$, then $K \dot{\div} b = K$,

(P4-Success) If $b \notin Cn(\emptyset)$, then $b \notin K \dot{\div} b$,

(P5-Extensionality) If $b \leftrightarrow c \in Cn(\emptyset)$, then $K \dot{\div} b = K \dot{\div} c$,

(P6-Recovery) $K \subseteq (K \dot{\div} b) + b$.⁵

Various representation theorems show that $\dot{\div}$ on K is a PMC operation if and only if $\dot{\div}$ satisfies P1-P6.

2.2 Contraction, Conservation and Entrenchment

In addition to the presentation of AGM belief dynamics based on PMC, Gärdenfors (1984, 1988), and Gärdenfors and Makinson (1984) showed that AGM belief dynamics could be interpreted in terms of the concept of *epistemic entrenchment*, or epistemic importance, that also provides an intuitively satisfying explication of contraction in terms of the POC. The basic intuition behind this interpretation of belief change is that we ought to give up those beliefs that are least entrenched. Gärdenfors and Makinson (1988) explain,

The epistemic entrenchment of a fact represents how important it is for problem solving or planning on the basis of the knowledge system and in this way determines the database *priority* of the fact (84).

In addition, using the older term epistemic importance, Gärdenfors (1984) explains that,

My main thesis is then that when we have to give up some of our beliefs we retain those with greatest epistemic importance (137).

This concept of epistemic entrenchment is given a more formal presentation as follows. Given that a and b are sentences, ' $a \leq b$ ' signifies that b is at least as entrenched as a , and ' $a < b$ ' signifies that b is more entrenched than a where $a < b = a \leq b$ and $\neg b \leq a$. Epistemic entrenchment is governed by the following postulates:

(E1-Transitivity) If $a \leq b$ and $b \leq c$, then $a \leq c$.

(E2-Dominance) If $a \vdash b$, then $a \leq b$.

(E3-Conjunctiveness) For any a and b , $a \leq (a \& b)$ or $b \leq (a \& b)$.

(E4-Minimality) If K is consistent, then $a \notin K$ if and only if $a \leq b$ for all b .

(E5-Maximality) If $b \leq a$ for all b , then $\vdash a$.

These postulates yield a comparative ordering on K of all sentences a, b, c, \dots of K in \mathcal{L} , and representation theorems show that \leq on K is an EEC (*epistemic entrenchment contraction*) if and only if \leq satisfies E1-E5.

Perhaps more interestingly Gärdenfors (1988) proved the following central theorem of the AGM theory:

(AGMT) A contraction function $\dot{\div}$ satisfies P1-P6 if and only if \leq satisfies E1-E5, where $b \leq a$ if and only if $b \notin K \dot{\div} (a \& b)$ for all a and b in \mathcal{L} .⁶

So, it should be obvious that the AGM concept of a contraction can be interpreted either in terms of the basic concept of minimal belief change (PMC) that satisfies P1-P6 or in terms of the basic concept of epistemic entrenchment (EEC) that satisfies E1-E5. Here I will focus on criticism of the AGM theory interpreted in terms of the concept of epistemic entrenchment, but, by AGMT and other related equivalence results, it should be obvious that those criticisms will apply *mutatis mutandis* to

the interpretation of the AGM theory in terms of the concept of partial meet contraction.

3. The Coherentist Interpretation of the AGM Theory

Having sketched out the AGM theory of belief revision we may now turn our attention to the critical appraisal of Gärdenfors' version of this theory interpreted in terms of coherentism, but first a brief discussion of epistemological coherentism seems desirable in order to see why one might be tempted to give a coherentist interpretation of the AGM theory. First and foremost, Gärdenfors appears to have been compelled to offer an epistemological interpretation of the AGM theory due to the general perception that the AGM theory was much too formal, and that, as a result, it had little or no relevance either to actual epistemic agents or to traditional epistemological concerns.⁷ In an effort to bring this influential formal theory into congruence with traditional epistemological views, and to avoid these accusations of excessive idealization and abstraction, Gärdenfors (1988, 1992) ultimately suggested that the AGM theory was most similar to traditional versions of epistemological coherentism like those defended by Laurence Bonjour, et al.⁸

The similarity between the AGM theory and epistemological coherentism is, of course, primarily due to the fact that the AGM theory, like coherentism, focuses on belief states defined by *DfBS*, and, as with coherentism, one of the minimal requirements (i.e. *DfBS(i)*) that an acceptable belief state must satisfy is that of logical consistency. As such, the AGM theory of belief dynamics seems naturally to admit of interpretation as a sort of dynamic version of traditional epistemological coherentism. Moreover, it is the avowed internalistic character of the AGM theory that, at bottom, suggests that the AGM theory is a coherence theory; i.e. on the AGM theory the epistemic status of beliefs is a wholly a matter of states internal to the epistemic agent.⁹

3.1 The Concept of Coherence

Anyone even mildly familiar with the history of the coherence theory is surely aware that the main problem plaguing coherence theories has been that of providing a sufficiently clear and precise account of the concept of coherence, the concept at the very heart of the coherence theory of knowledge and justification. For the purposes of examining the similarity between the AGM theory of belief revision and coherentism I will here characterize coherence in the manner suggested by Bonjour and Lewis, and which I think is the most plausible version of that theory in spite of many well-known criticisms.¹⁰

In Bonjour 1985 coherence is understood informally as follows: According to the coherence theory of empirical justification, as so far characterized, the epistemic justification of an empirical belief derives entirely from its coherence with the believer's overall system of empirical beliefs and not at all from any sort of factor outside the system (101).

So, the main points that we need to focus on here are, first, that the coherentist regards justification to be a matter of the coherence of a whole belief state, and, second, because it is a form of epistemological internalism, that this feature must be accounted for in terms of the beliefs that constitute such a state and their internal relations.¹¹ So it is clear that for coherentists of this stripe that the primary unit of epistemic interest is the belief state (i.e. a set of beliefs and relations among them represented as sentences standing in certain relations to one another) and that coherence is primarily an internally determined property of whole belief states. The justification of individual beliefs is then a secondary, derivative, concept. The obvious implication here, however, is that a coherent belief is, in some manner, a belief that is a member of a coherent belief system.

Let us now return to the obvious question that immediately arises in considering the acceptability of coherentism; the question of the specific nature of coherence. One of the primary tasks of Bonjour 1985 and 1999 was to make some headway towards an answer to this central question about epistemological coherentism. In his 1999 article Bonjour tells us that,

On an intuitive level, coherence is a matter of how the beliefs in a system of beliefs fit together or "dovetail" with each other, so as to constitute one unified and tightly structured whole. And it is clear that this fitting together depends on logical, inferential, and explanatory relations of many different sorts among the components of the system (123).

Expanding on this comment Bonjour notes that any adequate notion of coherence must satisfy the following desiderata: (a) it must require logical consistency of belief states, (b) it must involve inferential connectedness among the beliefs that constitute belief states of both deductive and probabilistic sorts, and (c) it must not exclusively make coherence a matter of explanatory connections among the beliefs of the belief state. But, this is by no means an especially clear account of coherence, and in order that we can evaluate this type of view what we would like is a much more clear definition of just what coherence is in terms of the representation of doxastic/epistemic states as sets of sentences.

Fortunately, Shogenji (1999, 2001) has recently specified some independent, but related, desiderata for any adequate definition of coherence that help to flesh out this admittedly sketchy epistemological notion. Shogenji (1999) says of such a definition that,

First, it should capture reasonably well what epistemologists ordinarily take coherence to be. This includes the idea that coherent beliefs 'hang together'; that coherence comes in degrees; and that coherence is symmetrical – i.e., the order of beliefs examined should not affect the assessment of their coherence. Second, the definition should be precise in that it provides a quantitative measure of coherence, thereby allowing us to compare diverse sets of beliefs in terms of their degrees of coherence (338).

Finally, as coherentism is clearly a version of internalism as noted in section 3.1, we can add that the property of coherence should turn out to be a supervenient property of belief states.¹² Combining these desiderata for the concept of coherence with those suggested by Bonjour we get the following set of adequacy conditions for the explication of the concept of coherence;

- (a) coherence must require logical consistency of belief states,
- (b) coherence must involve inferential connectedness among the beliefs that constitute belief states of both deductive and probabilistic sorts,
- (c) coherence must not exclusively make coherence a matter of explanatory connections among the beliefs of the belief state,
- (d) coherence must admit of degrees,
- (e) coherence must be symmetrical,
- (f) coherence must be a precise, quantitative, measure, and
- (g) coherence must be a supervenient property of belief states.

Given these adequacy conditions I suggest that, at least for the sake of the arguments to be presented here, we can accept Shogenji's 1999 proposal that we define the coherence of a set of sentences representing the epistemic state of an agent as follows:

$$(SC) C(b_1, \dots, b_n) = P(b_1 \& \dots \& b_n) / P(b_1) \times \dots \times P(b_n).^{13}$$

This measure is taken to represent the joint likelihood of the belief state and it requires logical and probabilistic consistency. Moreover, SC identifies various inferential connections (both deductive and inductive) among elements of belief states that can be defined in probabilistic terms, and, insofar as there are probabilistic relations that are non-explanatory, SC does not make coherence merely a matter of explanatory relations among beliefs in some *K*. Additionally, the measure *C*(·) seems naturally to admit of interpretation as a supervenient property of *K*, where *K* is (b_1, \dots, b_n) , that is completely determined by the (internal) inferential relations that obtain among the elements of *K*. So SC allows us to satisfy all of adequacy conditions (a)-(g) for the concept of coherence noted above, and it appears to be an eminently plausible construal of coherence. As such, I will assume, at very least, that coherence must satisfy (a)-(g), while noting that Shogenji's SC is one of the better explications of that concept. In any case, having now examined both the AGM theory of belief

revision and the basic concepts of epistemological coherentism we can now turn our critical attention to evaluating Gärdenfors' attempt to underwrite the former by appeal to the latter.

4. The Evidence Concerning Our Justifications

Coherence theories of knowledge and justification are, of course, old hat, but the use of such theories in support of the AGM theory of belief revision is novel. So, responding to the criticisms noted in previous sections, Peter Gärdenfors (1988, 1992) proposed that the AGM theory could be interpreted in terms of a more or less familiar version of the coherence theory of knowledge and justification. As we can now see this move was intuitively plausible and epistemically satisfying because of the basic similarity between the AGM method of belief representation and those of garden-variety coherentism. Specifically, both coherentism and the AGM theory employ the same type of representation of epistemic states and their internal relations, that epistemic states can be represented in some appropriate language by sets of logically consistent sentences characterized by internal inferential relations. However, were this all that Gärdenfors has to say about the relationship between the AGM theory and coherentism, then his claims concerning the similarity between AGM-style epistemic representation and coherentist-style epistemic representations might, for the most part, be regarded as unobjectionable.¹⁴ But, this recognition, of course, is not all that Gärdenfors has to say concerning the matter.

Taking a cue from Gilbert Harman (1986) and appealing to Ross and Anderson's (1982) compilation of studies concerning belief perseverance, Gärdenfors (1992) explicitly argues that empirical evidence shows that the coherentist interpretation of the justificatory status of our beliefs is likely to be true. This is indeed a more interesting response than merely noting similarities between the AGM theory and coherentism. In short, if the AGM theory is a dynamic version of coherentism and the evidence cited in Ross and Anderson is confirmatory with respect to coherentism, then we have some good, empirical, reasons to believe that the AGM theory is a literally true description of at least some aspects of human rationality. If this chain of reasoning is acceptable, then Gärdenfors can claim to have answered both the criticism that the AGM theory is too idealized and the criticism that AGM theory is out of tune with traditional epistemological concerns. In the sequel we shall see, however, that this is a chain of reasoning riddled with conceptual flaws.

4.1 *Belief Perseverance and the Evidence for Epistemic Coherence*

In effect, what Harman (1986) and Gärdenfors (1992) argue by appealing to Ross and Anderson (1982) is that empirical evidence shows that real epistemic agents do not store the justifications for their beliefs and that real epistemic agents do not give up beliefs when confronted with information that defeats the (apparent) justifications for beliefs that agents already hold. More importantly, they suggest that foundationalist epistemologies are committed to the view that a given belief is justified if and only if it is (currently) based in some way on another belief that serves as evidence for the given belief. As a result, if the justifying belief is subsequently defeated, then, according to the foundationalist, the agent must give up the belief(s) for which the justifying belief served as evidence. This is the substance of what we might call, in opposition to the POC, the principle of pedigree (POP). More explicitly the POP can be understood as:

(POP) When changing beliefs in response to new evidence, you should continue to believe only those beliefs that are justified in your new belief state.

But we might now ask specifically why Harman and Gärdenfors interpret the belief perseverance data as being evidentially supportive with respect to coherentism.

The main point that Harman and Gärdenfors attempt to make concerning the studies conducted by Ross, et al.,¹⁵ and which are summarized in Ross and Anderson 1982, is that these studies indicate that agent's actual beliefs are much more resilient than would be expected on the foundationalist model, which rests on the POP. Real epistemic agents seem to retain beliefs in spite of the fact that the (formative) justifications they have for such beliefs are in fact erroneous and that agents have been made aware of this; i.e. that their justifications have been defeated and that they know this. The studies by Ross and his colleagues show that it is true that beliefs are highly perseverant in this way, and Harman and Gärdenfors take this to be evidence that foundationalism is likely to be false. Thereby, they imply that coherentism is more likely to be true, as coherentism does not require that beliefs must be justified in virtue of other undefeated beliefs.

Recall, that for coherentists justification is a supervenient property of whole belief states, and so Harman and Gärdenfors seem to be arguing that the belief perseverance results are *best explained* by supposing that coherentism is true,¹⁶ and that foundationalism identifies justifications ultimately with basic beliefs and so does not explain the data well, if at all. Gärdenfors tells us that,

According to the coherence theory, beliefs do not usually require any justification; the beliefs are justified just as they are (1992, 380).

Of course, this is because, for coherentists justification is a holistic property of belief states determined by their manner in which the component beliefs of belief states mutually support one another. As such, on the coherence view, the justification of a belief may resist the defeating of a belief that, on a foundationalist view, would require a rational agent to cede that same belief. So, according to Harman and Gärdenfors, the coherentist is in this way supposed to be able to explain the phenomenon of belief perseverance in a considerably more plausible way than the foundationalist is able to do.

Moreover, Harman (1986) appeals to a principle of informational economy, much like the POC, in attempting to provide a more positive explanation of this kind of belief perseverance.¹⁷ He simply argues that it is highly impractical for epistemic agents to store the original justifications for their beliefs given their finite and highly limited storage capacities and that such information is a valuable epistemic resource. We see that Gärdenfors essentially concurs when he tells us 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 (1992, 381).

Retention of information, then, in spite of evidential undermining, is an epistemic virtue for Harman and Gärdenfors, and coherentism best explains why this is so. The real epistemic agents in the belief perseverance studies then must be regarded as being rational in terms of typical versions of coherentism as they, in point of fact, tend to cling tenaciously to beliefs in the manner the POC suggests that they should.

4.2 *The Empirical Misinterpretation of Belief Perseverance Data*

In this section I will argue that the Harman and Gärdenfors' interpretation of the belief perseverance experiments as favoring coherentism is deeply flawed. Specifically, I will argue their claim that the empirical evidence presented by Ross, et al is best explained by coherentism is false. If this criticism is correct, then Gärdenfors' attempt to provide empirical support to the AGM theory via its relationship with coherentism, and thereby to deflate the accusations of excessive idealization and abstraction of that theory fails.

Recall that Harman and Gärdenfors both present the evidence summarized in Ross and Anderson 1982 as providing empirical support for coherentism. Moreover, they also argue that foundationalist views cannot explain this phenomenon as easily, if at all. But, this claim is simply false and, in fact, does not accord well with the conclusions that the empirical researchers themselves make.¹⁸ Consider Lord, Ross, and Lepper's conclusion that,

If our study demonstrates anything, it surely demonstrates that social scientists can not expect rationality, enlightenment, and consensus about policy to emerge from their attempts to furnish "objective" data about burning social issues (1979, 2108).

Similarly, Ross, Lepper, and Hubbard explain that,

The experiments reported in the present article were designed to demonstrate that once formed, impressions are remarkably perseverant and unresponsive to new input, even when such input logically negates the original basis for the impressions...As a result, subsequent challenges to that evidence, and hence to the impression it fosters, will have surprisingly little impact-far less impact than would be demanded by any logical or rational impression-formation model (1975, 880).

An obvious implication is then that the subjects' behavior could well be regarded as irrational. But elsewhere, Anderson, Lepper and Ross are hesitant about drawing this strong conclusion about human irrationality from the perseverance experiments. Although they explain that the experiments,

[...] offer further evidence for the basic hypothesis that people often cling to their beliefs to a considerably greater extent than is normatively warranted (1045),

they are also careful to add that,

Our argument, then, is not that the mechanisms underlying belief perseverance are inherently rational or inevitably dysfunctional. Rather, we are suggesting that these mechanisms, like many other processes underlying human inference, may lead in certain contexts to a normatively unwarranted judgment, belief, and behavior (1980, 1047).

So, wisely, Ross and his colleagues seem to be arguing neither that their experiments show that humans are rational nor that they show that they are irrational. Rather, they simply report the existence of the phenomenon of belief perseverance and offer some speculation about the mechanisms by which it works.

Harman and Gärdenfors however claim that the empirical data indicating the belief perseverance phenomenon turns out to be clearly *indicative* of rational behavior. This is no surprise given that they accept a form of the POC in conjunction with coherentism, and so the behaviors of the subjects of the experiments in question turn out to be exemplary cases of rational behavior on their views. In accord with the POC the experimental subjects do not give up beliefs that, for foundationalists, would rationally have to have been given up, and so long as those beliefs are coherent, then, on Harman and Gärdenfors' views these subjects acted rationally. Gärdenfors (1992) elaborates;

So do these empirical findings mean that people are irrational? Shouldn't they try to keep track of their beliefs? I think not. The main reason is that it is intellectually extremely costly to keep track of the

sources of beliefs, and the benefits are, by far, outweighed by the costs. A principle of intellectual economy would entail that it is rational to neglect the pedigree of one's beliefs. To be sure, we will sometime hold onto unjustified beliefs, but the erroneous decisions caused by this negligence must be weighed against the cost of remembering all reasons for one's beliefs. The balance will certainly be in favor of forgetting reasons. After all, it is not very often that a justification for a belief is actually withdrawn and, as long as we do not introduce new beliefs without justification, the vast majority of our beliefs will remain justified (383).

Surely, this is a very odd interpretation of that data presented by Ross and his colleagues. As we have already seen, Ross and his colleagues interpret the data neither as being indicative of rational epistemic behavior, nor as being evidence for the claim that the experimental subjects acted in irrational ways. This seems far more plausible and much more cautious, and, at very least, shows that there is an alternative explanation of that data available to the foundationalist. One could easily argue that the data is better explained by the foundationalist account of justification in conjunction with the POP, but that humans often fail to meet such standards.¹⁹ The foundationalist's explanation of the belief perseverance data can also be made much more plausible by making the natural assumption that justifying beliefs are, by and large, *implicit* beliefs. As such, it would be no surprise that the experimental subjects failed to cede beliefs that, given POP, they should have ceded.

Gärdenfors' primary argument in the passage above is that the principle of intellectual economy, the POC, entails that we should not keep track of the justifications of our beliefs, that we should neglect the pedigree of our beliefs, thus rejecting the POP. In effect, what he argues is that if the POC is true, then coherentism is a better explanation of the empirical data than foundationalism. However, this does not provide sufficient reason to draw the conclusion that coherentism *really* is the best explanation of the perseverance data unless there is some independent motivation for accepting the POC. It simply begs the question against the foundationalist who accepts the POP. Gärdenfors' only claim in support of the POC is that although accepting the POC will sometimes result in our retaining unjustified beliefs, the cost of remembering our justifications outweighs this risk because beliefs are not often withdrawn. What is the justification for this claim? Why should we claim that it is rational to retain beliefs in a way that maximizes tenacity? As far as I can tell, Gärdenfors offers no justification at all, and simply begs the question against those who would accept the POP. So, absent any substantive justification for the POC, it is just as plausible to believe that one ought to accept foundationalism and the doctrine that belief is a matter of pedigree. This implies that the subjects ought to

reject beliefs when the justifications for such beliefs are defeated, and so they simply acted irrationally when they did not.

At very least then, it seems that the perseverance data do not *clearly* favor the coherentist theory over the foundationalist theory. That would only be the case if one were to accept the quite dubious assumption that the experimental subjects were acting rationally in those experiments by accepting the POC. However, as we have seen, Gärdenfors does not give us any good reasons to accept the POC, and Gärdenfors' own acceptance of the POC is merely a matter of his unsupported intuition to this effect. Underdetermination considerations then imply that Harman and Gärdenfors' attempt to exploit the results of Ross', et al, experiments as providing evidence for coherentism via inference to the best explanation utterly fail. The behavior of the subjects is equally well explained both as a case of the irrational behavior of foundationalist agents and as a case of the rational behavior of coherentist agents. So Harman and Gärdenfors' explanation is, at best, no better than this alternative explanation, and since it is relatively uncontroversial that confirmation and explanation are differential matters, the belief perseverance data does not support the coherence theory in the manner they suggest.²⁰

5. Justification as Entrenchment

In spite of the errors made in interpreting the significance of the results collected in Ross and Anderson 1982 and in attempting to employ those results in support of coherentism, Gärdenfors has attempted to develop the coherence version of the AGM theory, and, in doing so, has suggested that the traditional concept of justification can be effectively replaced with the concept of epistemic entrenchment as defined in section 2.2. In this section I will present what I take to be three rather serious conceptual criticisms of this suggestion. Collectively, I will argue that these criticisms undermine Gärdenfors' attempt to replace the concept of justification with that of entrenchment irrespective of the fact that coherentism is itself not confirmed by Ross and Anderson's (1982) collection of empirical data. As such, his attempt to bring the AGM theory into touch with traditional epistemological concerns by showing that the AGM theory is a version of coherentism fails.

5.1 Degrees of Justification and Comparisons of Entrenchment

As we can now see, the main project that Gärdenfors appears to be undertaking in an effort to ally the AGM theory and coherentism is an attempt to replace the traditional concept of justification with the con-

cept of entrenchment. In addressing broadly foundationalist criticisms of the AGM theory Gärdenfors (1992) explains,

It is now, finally, time to return to the challenge from the foundations theory that models of epistemic states based on the coherence theory cannot express that some beliefs are justifications or reasons for other beliefs. I want to show that if we take into account the information provided by the ordering of epistemic entrenchment, the belief sets of the AGM model can, at least to some extent, handle this problem (390).

A bit later he claims that,

[...]if we consider the extra information provided by the ordering of epistemic entrenchment, then this information suffices to account for much of what is required by the foundations theory of belief revision. A belief set in itself does not contain any justifications of the beliefs included. However, if an ordering of epistemic entrenchment is added to such a belief set, than at least some justifications for the beliefs can be reconstructed from this ordering (392).

So the concept of entrenchment, to some degree, is alleged to be able to fulfill the same function as that of justification, and it seems, essentially, that this project is one that is best understood as a case of replacing a loose concept with a more formal one. As such, Gärdenfors' project, at very least in this specific respect, is a case of the explication of a concept.²¹ However, as with all good explications, the explicandum must be such that it, at least, fulfills the same function as the explicandum so that the former can be used in place of the latter. Moreover, the explicatum is supposed to be both more precise and more well defined than the explicandum. I will argue here that there are several, very compelling, reasons to suspect that Gärdenfors' replacement project does not fulfill these requirements for proper conceptual explication.

To begin, I take it to be trivially true that the traditional (theory-neutral) concept of justification admits of continuous numerical degrees, and that justification (even of the coherentist sort) is measurable in terms of probabilities in just the way, for example, that SC attempts to provide. The function of the concept of justification seems to be one of allowing us to not only compare the justifications of beliefs, but also to assess the degree of justification that a belief has.²² This, in turn, explains the deeply normative function of the concept of justification. The concept of entrenchment, as presented by Gärdenfors and as defined in 2.2, is, however, merely a comparative concept, and, as such, does not admit of continuous numerical degrees.²³ So, it does not satisfy adequacy conditions (d) and (f) from section 3.1, and, hence, it certainly cannot be an adequate explication of *coherentist* justification. These considerations indicate that justification is, in fact, a *more precise concept* than entrenchment, and that entrenchment cannot be an acceptable explication even of the basic, theory-neutral, concept of justification as in

adequate explications the explicatum must be more precise than the explicandum.

5.2 The Normative Function of Justification and the Facts About Entrenchment

As noted in section 5.1, justification is a normative concept and Gärdenfors argues that entrenchment can serve as a replacement for that concept, but it is not clear that entrenchment is in any way a normative concept. As such, it is not clear that entrenchment can serve the same basic function as the concept of justification, even if it could be made precise in such a way that it would admit of numerical degrees. Gärdenfors explains the function of entrenchment by telling us that,

Not all sentences that are believed to be true are of equal value for planning or problem-solving purposes, but certain pieces of knowledge about the world are more important than others when planning future actions, conducting science, or reasoning in general (1992, 387).

More crucially, he tells us that the comparative concept of entrenchment is defined only in relation to a given *K* that it is a comparative measure of informational value or epistemic importance (1992, 387). However, elsewhere, he tells us that there is no accepted, or even well understood, account of the relative informational value of beliefs (1988, 91-94). As Gärdenfors' 1988 discussion indicates, it is not at all clear that information is even an objective property, let alone a normative property. So the basic concept of the informational value of beliefs that is supposed to yield orderings of epistemic entrenchment relative to a given *K* is a matter on which we are left in the dark, and which might well turn out to wholly subjective and purely descriptive.²⁴ A given agent might tenaciously hold onto a given belief, say that God exists, in a way that makes it almost maximally entrenched, but why should we accept that this is what the agent should do simply because that belief is, *in fact*, deeply entrenched relative to his belief state? Such a belief might well be epistemically useful, or important, to that agent, but this could be true of any belief and for any agent. This is especially troubling as there will be many possible orderings of the beliefs that constitute a belief state in terms of their entrenchment. Another agent might well have the very same belief state, but an entirely different ordering of entrenchment defined on that belief state.

We cannot then obviously treat entrenchment as normative and objective in the manner that justification, especially of the coherentist sort, is alleged to be. This is because coherentists are internalists, and it is rather uncontroversial that adherents of this epistemological view are committed to the principle that epistemic agents that are identical *qua*

their belief state, must be identical *qua* their justificational state.²⁵ This does not seem to be true of entrenchment, and unless we are provided with an objective account of the informational value of beliefs we have no reason to suspect that epistemic agents that are identical *qua* their belief state, must be identical *qua* their entrenchment ordering. As a matter of fact we have not been provided with such an account, and I am frankly dubious that there is such an account to be had. More importantly, as we have no extant objective account of the concept of informativeness, it does not seem that entrenchment can be seriously regarded as a well-defined concept. As such, it cannot fulfill the functional role of the concept of justification, either of the coherence sort or otherwise, and so cannot be an adequate explication of justification. In short, entrenchment is neither well defined, nor is it functionally equivalent to coherentist justification.

5.3 Why Coherentist Justification cannot be Entrenchment: The Incoherence of Entrenchment

The coherence theory of justification of the sort defended by Bonjour holds that justification is a property of belief states, and so given this theory one's justification is a holistic matter. Moreover, on this view, justification is coherence and coherence is a property of belief states, not of individual beliefs (except perhaps in a derivative way). Entrenchment, on the other hand, is explicitly a property of beliefs, albeit a relational property that holds (minimally) between two beliefs and a belief state. This suggests that there are serious logical asymmetries between the two concepts, and that Gärdenfors' attempt to undergird the AGM theory by appeal to coherentism is purely chimerical. That *b* is more entrenched than *a* relative to *K* is supposed to show that *b* is more justified than *a*, because *b* will be less easily ceded than *a*. But, how can this be a matter of the coherence of *K* as described in earlier sections?²⁶ They are totally different types of properties.

So, even if the experimental evidence from Ross and Anderson 1982 could be shown to be confirmatory in the way Gärdenfors erroneously suggests it is and the concept of entrenchment could be made both more precise and well defined, it cannot be a suitable replacement for the concept of coherentist justification. Entrenchment just isn't a property of belief states as wholes, and, in accord with adequacy condition (g) from section 3.1, coherence is. To put it baldly, the properties of entrenchment and of justification are not of the same logical type. The latter is a unary predicate, and the former is a complex relational predicate.²⁷ This simple fact blocks Gärdenfors' attempt to appeal to epistemological coherentism as way of bringing the AGM theory into congruence with traditional epistemological concerns.

6. Rational Belief, Idealization, and the Methodology of Epistemology

Having now seen both the failure of Gärdenfors' attempt to support the AGM theory by appeal to coherentism and the belief perseverance studies, and the virtual incoherence of his suggestion that the concept of entrenchment can replace that of justification in such a theory, it is apparent that the original perception of the deficiencies of the AGM theory are quite accurate. The AGM theory is not plausibly interpreted as a version of coherentism, and neither the AGM theory nor coherentism is supported by the belief perseverance data.

What lessons can we learn from this failure? I think that there are two main lessons. First, we ought to be extremely careful about drawing epistemological conclusions from empirical studies. Even naturalists should observe this dictum as such empirical results are subject to typical problems of underdetermination. This is further complicated by the apparent problems of drawing normative conclusions from empirical results, and unless one is willing to accept the most radical, and rather implausible, form of descriptive epistemological naturalism we ought to be very careful in drawing any deep methodological conclusions from such studies.

Second, we ought to be extremely careful about trying to connect formal, highly idealized and abstract, theories of rationality with traditional epistemological theories. Often, due to the very imprecise nature of many component concepts in traditional epistemological theories, the fit will be, at best, poor. What are not needed are further *ex post facto* attempts to link *a priori* generated logical formalisms with *a priori* generated epistemologies and bodies of empirical evidence. Such approaches are the worst kinds of examples of 'top-down' approaches to modeling and problem solving. In such approaches utterly speculative (*a priori*) theories, masquerading as traditional epistemological theories, are forced onto empirical results without regard for the possibility that the formal model itself might simply be descriptively false, and the feeble response that epistemology is an *a priori* discipline does nothing to dispel such problems. What are needed are more serious and careful attempts to construct normative theories of rationality beginning with, and under the constraint of, empirical studies concerning human epistemic states and their dynamics; theory should follow, and harmonize with, experimental evidence. Surely idealizations are permissible in epistemology, but they must be suggested by empirical study and not introduced without any such objective, empirical, guidance.²⁸

Notes

¹ Of course, it is true that other criticisms of the AGM theory have been presented elsewhere. For example, Tennant 1994 and 1997 focus on primarily technical objections to the AGM postulate of recovery (P6 in section 2). More fundamental and specifically epistemological criticisms have been raised in Hansson and Olsson 1999 and in Pollock and Gillies 2000.

² Hansson 1999 is an extensive, and considerably more detailed, introduction. Gärdenfors 1988 is the canonical presentation of the theory, however. In what follows I will remain faithful to both, but will also make use of elements of Gärdenfors and Makinson 1984, Olsson 1997, Hansson and Olsson 1999, and Rott 2000.

³ In point of fact, the AGM theory really only holds that there are two dynamical operations on belief states as revision is defined in terms of expansion and contraction.

⁴ This principle will be further investigated in section 4.1.

⁵ This presentation of the AGM contraction postulates follows Gärdenfors 1988 and Hansson and Olsson 2000 most closely with only minor notational variations to yield consistency of formalism.

⁶ Rott 1991 also includes a related proof that the concepts of epistemic entrenchment contraction and partial meet contraction are strictly equivalent

⁷ This suspicion appears to be justified from looking at Doyle 1992, and, especially, the introduction to Gärdenfors 1988, and Gärdenfors 1992. This criticism has also recently been stressed in Pollock and Gillies 2000.

⁸ See, for example, BonJour 1985. Also, see the discussion in Lewis 1946.

⁹ See Gärdenfors 1988, 18-20.

¹⁰ I omit discussion of coherentism in the sense defended by Lehrer in his 1974, 1989, 1990a and 1990b and this is because coherence in the sense employed by AGM theorists is a property of belief states and not a relation between a belief and a belief state as in Lehrer's version. Moreover, it is important to note that BonJour (1999), however, has explicitly rejected the theory from BonJour 1985 due to the battery of criticisms leveled at it over time. Many of these criticisms can be found in Bender 1989, but Haack 1993 also raises serious, related, criticisms.

¹¹ See Bender 1989, Chisholm 1988, Greco 1990, Kornblith 1988, and Feldman and Conee 2001 for discussion of internalism and its relationship to coherentism.

¹² Feldman and Conee 2001 discuss this aspect of internalism and claim that all internalists are committed to the view that, "the justificatory status of a person's doxastic attitudes strongly supervenes on the person's occurrent and dispositional mental states, events, and conditions (2)."

¹³ Of course this proposal may not be wholly acceptable as Akiba 2000 and Olsson 2001 suggest. In any case, Shogenji has addressed Akiba's criticisms in Shogenji 2001a. For a more thorough discussion of the role of coherence in justification see Shogenji 2001b.

¹⁴ This is not, however, actually the case as we shall see in section 5.

¹⁵ See, especially, Ross, Lepper, Hubbard 1975, Lord, Ross, and Lepper 1979, and Anderson, Lepper, and Ross 1980.

¹⁶ This way of construing Harman's argument seem especially appropriate as he has famously argued that all inductive arguments are cases of inference to the best explanation in Harman 1965.

¹⁷ See Harman 1986, chapter 4.

¹⁸ See, for example, the concluding sections of Ross and Anderson 1982.

¹⁹ Stein 1996 provides an especially good example of why we should be very careful about how we regard empirical studies concerning human rationality, and how we relate such studies to epistemological theories.

²⁰ See Erwin and Siegel 1989 for an extensive discussion of differential confirmation, including confirmation based on inference to the best explanation. Also, see Lipton 1991 for related discussion.

²¹ See Carnap 1956 on the concept of explication.

²² This basic view is shared by all epistemologists who accept probability as a measure of the degree of justification, and few deny this basic view.

²³ See Gärdenfors 1984 and 1988.

²⁴ See, for example, Kreider 2001.

²⁵ See Feldman and Conee 2001, and adequacy condition (g) from section 3.1.

²⁶ That entrenchment cannot be coherentist justification can be further illustrated by simply noting that it seems perfectly possible that there could be belief states on which an ordering of epistemic entrenchment is defined, but which are clearly incoherent, at least as coherence is here understood. Consider a belief state constituted by three probabilistically independent propositions. Surely, there can be comparative differences in the epistemic importance or informational value of these propositions, and, hence, they will be ordered in terms of entrenchment, but they cannot be coherent as there are, *ex hypothesi*, no inferential relations at all between them.

²⁷ Depending on whether informativeness, or epistemic importance, turns out to be objective or subjective will determine whether entrenchment is a triadic or quadratic relation. If it is subjective, then it will be quadratic; i.e. a is more entrenched than b relative to K for epistemic agent S . If it is objective then it will be triadic; i.e. a is more entrenched than b relative to K . Either way, it is not logically of the same type as coherence.

²⁸ This point about the empirical constraints on idealization in general is defended in Shaffer 2000 and Shaffer 2001. Exemplary cases of this attitude in epistemology are to be found in Cherniak 1986 and Stein 1996.

Bibliography

- Akiba, K. (2000) "Shogenji's Probabilistic Measure of Coherence is Incoherent," *Analysis* 60: 356-359.
- Aichourrón, C., P. Gärdenfors and D. Makinson (1985) "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions," *The Journal of Symbolic Logic* 50: 510-530.
- Anderson, C., M. Lepper, and L. Ross (1980) "Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information," *Journal of Personality and Social Psychology* 39: 1037-1049.
- Bender, J. (1989) *The Current State of the Coherence Theory*. Dordrecht: Kluwer.
- BonJour, L. (1985) *The Structure of Empirical Knowledge*. Cambridge: Harvard University Press.
- _____. (1999) "The Dialectic of Coherentism and Foundationalism," in J. Greco and E. Sosa, *The Blackwell Guide to Epistemology*. Malden: Blackwell: 117-142.
- Cherniak, C. (1986) *Minimal Rationality*. Cambridge: MIT Press.
- Chisholm, R. (1988) "The Indispensability of Internal Justification," *Synthese* 74: 285-296.
- Doyle, J. (1992) "Reason Maintenance and Belief Revision," in P. Gärdenfors, *Belief Revision*. Cambridge: Cambridge University Press.
- Erwin, E. and H. Siegel (1989) "Is Confirmation Differential?" *British Journal for the Philosophy of Science* 40: 105-119.
- Feldman, R. and E. Conee (2001) "Internalism Defended," *American Philosophical Quarterly* 38: 1-18.
- Gärdenfors, P. (1984) "Epistemic Importance and Minimal Changes of Belief," *Australasian Journal of Philosophy* 62: 136-157.
- _____. (1988) *Knowledge in Flux*. Cambridge: MIT Press.
- _____. (1992) "The Dynamics of Belief Systems: Foundations Versus Coherence Theories," in C. Bicchieri and M. L. Dalla Chiara, *Knowledge, Belief and Strategic Interaction*. Cambridge: Cambridge University Press: 377-396.
- Gärdenfors, P. and D. Makinson (1988) "Revisions of Knowledge Systems Using Epistemic Entrenchment," in M. Vardi (ed.), *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*. Morgan Kaufmann Publishers: 84-95.
- Greco, J. (1990) "Internalism and Epistemically Responsible Belief," *Synthese* 85: 245-277.
- Haack, S. (1993) *Evidence and Inquiry*. Malden: Blackwell.
- Hansson, S. O. (1999) *A Textbook of Belief Dynamics*. Dordrecht: Kluwer.
- Hansson, S. O. and E. Olsson (1999) "Providing Foundations for Coherentism," *Erkenntnis* 51: 243-265.
- Harman, G. (1965) "Inference to the Best Explanation," *The Philosophical Review*, 74, 88-95.
- _____. (1986) *Change In View*. Cambridge: MIT Press.
- Kornblith, H. (1988) "How Internal Can You Get?" *Synthese* 74: 313-327.
- Kreider, A. (2001) *On the Likelihood of Finding a Satisfactory Naturalistic Account of Mental Representation*. Ph.D. dissertation. Miami: The University of Miami.
- Lehrer, K. (1974) *Knowledge*. Oxford: Oxford University Press.
- _____. (1989) "Coherence and Truth Connection," J. Bender (1989): 253-275.
- _____. (1990a) *Theory of Knowledge*. Boulder: Westview.
- _____. (1990b) *Metamind*. Oxford: Clarendon Press.
- Lewis, C. I. (1946) *An Analysis of Knowledge and Valuation*. LaSalle: Open Court.
- Lipton, P. (1991) *Inference to the Best Explanation*. New York: Routledge.
- Lord, C., L. Ross, and M. Lepper (1979) "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology* 37: 2098-2109.

Reflective Equilibrium and the Justification of Inferential Principles

William Max Knorpp
James Madison University

ABSTRACT: *In Fact, Fiction, and Forecast (1954), Nelson Goodman suggested an account of the justification of principles of inference. This method was later adopted by John Rawls as an account of the justification of moral principles, and called "reflective equilibrium." The reflective equilibrium account of the justification of general (normative) principles is probably the most influential method in contemporary analytic philosophy, and is generally thought to be a species of coherentism. In this paper I identify some crucial ambiguities in Goodman's statement of his view. Once these ambiguities are recognized, it becomes clear that there is a plausible but uninteresting methodological interpretation of reflective equilibrium, and an interesting but implausible "critical" interpretation. I argue that Goodman's account fails because it cannot make sense of the normativity associated with logic – that is, because it degenerates into logical nihilism. I also suggest that this is the reason that justificatory relativist views generally fail.*

1. Introduction¹

In *Fact, Fiction, and Forecast*, Nelson Goodman suggested a way of justifying principles of inference (both deductive and non-deductive) which has been called "reflective equilibrium" and "justification by balance".² There is no more influential account of the justification of general principles in contemporary Anglo-American philosophy. Reflective equilibrium accounts are most often taken to be coherence accounts, though many have denied that this is the correct characterization of the view. I will examine Goodman's reflective equilibrium account of validity (GREAV for short), note some ambiguities in Goodman's statement of the view, argue that because of these ambiguities it is not clear that the view is a coherence view, suggest that the most important interpretation of the view is relativistic, argue that the standard counter-example strategy of refuting the view is sub-optimal, and suggest a more difficult (but more direct and philosophically fruitful) way of refuting the view (a way that explains *why* the view fails). This strategy also throws some light on the nature of relativism generally, and the reasons for the failure of relativistic

Olsson, E. (1997) *Coherence*. Ph.D. dissertation. Uppsala: University of Uppsala.

_____. (2001) "Why Coherence is Not Truth-Conducive," *Analysis* 61: 236-241.

Pollock, J. and A. Gillies (2000) "Belief Revision and Epistemology," *Synthese* 122: 69-92.

Ross, L. and C. Anderson (1982) "Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments," in D. Kahneman, P. Slovic and A. Tversky (eds.) *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press: 129-152.

Ross, L., M. Lepper, and M. Hubbard (1975) "Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm," *Journal of Personality and Social Psychology* 32: 880-892.

Rott, H. (1991) "Two Methods of Constructing Contractions and Revisions of Knowledge Systems," *Journal of Philosophical Logic* 20: 149-173.

_____. (2000) "Two Dogmas of Belief Revision," *The Journal of Philosophy* 97: 503-522.

Shaffer, M. (2000) *Idealization and Empirical Testing*. Ph.D. dissertation. Miami: University of Miami.

_____. (2001) "Bayesian Confirmation of Theories that Incorporate Idealizations," *Philosophy of Science* 68: 36-52.

Shogenji, T. (1999) "Is Coherence Truth Conducive," *Analysis* 59: 338-345.

_____. (2001a) "Reply to Akiba on the Probabilistic Measure of Coherence," *Analysis* 61: 147-150.

_____. (2001b) "The Role of Coherence in Epistemic Justification," *Australasian Journal of Philosophy* 79: 90-106.

Stein, E. (1996) *Without Good Reason*. Oxford: Clarendon Press.

Tennant, N. (1994) "Changing the Theory of Belief Change: Towards a Computational Approach," *British Journal for the Philosophy of Science* 45: 865-897.

_____. (1997) "Changing the Theory of Theory Change: Reply to My Critics," *British Journal for the Philosophy of Science* 48: 569-586.