HARMAN, NEGATIVE COHERENTISM, AND THE PROBLEM OF ONGOING JUSTIFICATION

THOMAS D. SENOR

I. Introduction

Epistemologists typically assume that theories of epistemic justification require only a single analysandum, viz., "S is justified in believing that P at t." I believe this assumption to be mistaken. Its error lies in the failure to recognize, and pay proper respect to, the important distinction between 'coming to believe' and 'continuing to believe' a proposition.

There is no obvious reason to suppose that the conditions necessary and sufficient for justifiably coming to believe that "Abe Lincoln was assassinated at the theater" are identical to those necessary and sufficient for continuing to believe it. Hence, epistemologists would be well advised to distinguish what we might call 'initial justification' from 'ongoing justification.' A philosophical analysis of the former would specify the conditions under which one is justified in coming to believe P, while an analysis of the latter would tell us when one is justified in continuing to believe it. Of course, it might turn out that these sets of conditions are identical; but there is no guarantee of that, and on reflection, it seems unlikely.

The labels 'initial justification' and 'ongoing justification' are a bit misleading. They suggest that the dichotomy is between (i) becoming justified in holding a belief and (ii) continuing to be justified in a belief, rather than between (iii) being justified in forming a belief and (iv) being justified in continuing to hold it; and it is the latter distinction that interests me. One can easily see that the two distinctions are indeed distinct by thinking of a case in which one unjustifiably forms a belief and later gains possession of evidence

sufficient to justify the belief she already has. If her belief is now grounded or sustained by the good evidence, her continued belief will become justified. So a complete theory of what I am calling 'ongoing justification' will have to have two parts: one specifying how an ongoing belief can become justified, and another specifying the conditions under which an ongoing belief that has become justified in the past continues to be justified.

In this essay, I wish to have a look at a negative coherentist account of ongoing justification.[1] My reason for choosing this theory is twofold: first, negative coherentism has a certain prima facie plausibility as a theory of ongoing justification that it lacks as a theory of initial justification; and second, it has recently been defended by Gilbert Harman in his book *Change in View*.[2] This second point makes negative coherentism of particular interest for me because Harman is one of the few epistemologists to consider the problem of ongoing justification.[3] Hence, in what follows I shall expound Harman's position and the interesting argument that he gives for it, and then try to demonstrate that his case for negative coherentism fails and that there are good reasons for rejecting it as a theory of ongoing justification.

II. Preliminary Distinctions

Before launching into my discussion of Harman, I believe it is important to clarify the primary term of epistemic appraisal, viz., justification. There are two significantly different concepts, or families of concepts, of epistemic justification. The first of these is intimately connected to knowledge; in fact, this variety of justification can be identified with whatever it is that converts true belief into knowledge (for our purposes we can ignore Gettier-type problems). Let's call this concept of justification 'K-justification.' The second concept of epistemic justification is what William Alston, Alvin Plantinga, and others call 'deontological' or 'deontic' (call it 'D-justification'): one is D-justified in believing that p iff one has done one's epistemic duty[4] regarding believing that p.[5] In other words, being D-justified in believing that P is to believe P responsibly; to be un(D-)justified is to be epistemically blameworthy.

Clearly, it is this latter sense of justification that has close parallels with the moral justification of action.

One should not rule out the possibility that D-justification is necessary for K-justification; so we shouldn't necessarily think of the two notions as unrelated. The point is simply that they are distinct. Although the K/D-justification dichotomy has frequently been overlooked, it takes only a moment to see that these two concepts of justified belief are not identical. A child might believe that her mother is going to win the mayoral race because her overly-optimistic father insists that her mom 'can't lose.' But suppose that the mother stands very little chance of succeeding. In that event, even if it should turn out that the girl's belief is true, we will not conclude that she knew her mother would win, but only that she was lucky. Be that as it may, the child has not violated any epistemic duty in believing as she does. Hence, she is justified in the second sense of the term but not in the first; so the two senses are not identical.

There is a lot more that could be said about each sort of justification, but for reasons that will become clear later, I will content myself with discussing only the deontological variety. As I have said above, the deontological sense of justification has as its essence epistemic duty fulfillment, responsibility, and praise or blame. This variety of epistemic appraisal is common. Often when we say of a person that she is unjustified, we mean to imply epistemic culpability, that she is to be censured for believing as she does. This aspect of D-justification is very important since it implies that it is appropriate to treat the evaluated subject as responsible for her belief. To be responsible for an action in a way that makes praise or blame appropriate, an agent's performance of that action must be 'up to her.' That is, if we accept an epistemic equivalent of the "'ought' implies 'can' "principle, we will have to grant that "S is un(D-)justified in believing that P" entails that "S could have refrained from believing P." That is, S must be able to exercise a certain amount of voluntary control vis-a-vis believing that P.

But what sort of voluntary control over belief is required if one is to be an appropriate target of deontological evaluation? Not surprisingly, the time at which an agent must exercise the sort of control necessary for epistemic evaluation is different in cases of

coming to believe than it is for continued belief. If S comes to believe that P at t, and at t+1 we consider whether S was D-justified in coming to believe that P at t, we are presuming that S had voluntary control at t, the time at which the belief was formed. However, if our query at t+1 concerns the justificatory status of S's continued belief at t+1, then the requisite control is that which is exercised at t+1. Yet we still haven't clarified the type of control necessary for epistemic responsibility.

Epistemologists frequently distinguish between direct and indirect voluntary control, although there isn't any canonical way of marking this dichotomy. Cutting to the chase, we can specify the crucial necessary condition of direct voluntary control for continued belief as follows: S has direct voluntary control at t over her continued belief that P only if, at t, S can bring it about via a simple act of her will that she ceases to believe that P or that she continue in her belief. On an analogous construal of direct voluntary control for states of affairs involving action, I now have such control over the state of affairs of my being seated. For I can immediately bring it about that I am not seated, or if I chose to remain as I am, I can bring it about that I am seated. To have indirect voluntary control over a state of affairs A is to be able to bring about a state of affairs A* via a simple act of will such that the obtaining of A* tends to (i.e., will generally) lead to the obtaining of A. For example, while I cannot become a better writer by a simple act of the will, there are basic actions that I can perform which will tend to bring about the state of affairs of my being a better writer. So if I move my arm and hand in a certain way (a basic action) I'll bring it about that I grab my copy of White and Strunk, and performing other basic acts will lead to my reading it and that will lead to the improvement of my prose. It is important to keep examples like that involving sitting and writing in clear view; for despite appearances, the distinction that really matters here has less to do with causal immediacy than with temporal immediacy. An agent who can only indirectly bring about A, but who can bring it about surely via a very quick series of causally related events has what is crucial for direct control in a way that one who can bring about A via a simple act of will involving a significant temporal gap between the act and its desired effect does not. I state

the conditions of direct and indirect control as I do only because it is easier to give a moderately clear account of causal immediacy than to give a similar account of temporal immediacy.

Let's briefly take stock. I have been engaged in this discussion of direct and indirect control in an effort to make clear the options for one who claims that we have control over our believings. And anyone who finds the concept of D-justification to have application will find herself committed to the claim that we possess at least some mode of voluntary control over some of our beliefs.

It is time to turn our attention to Harman's theory. In time we will see that the discussion of this section is necessary to set the issue out properly.

III. Harman's Position

Let's begin by looking at the theory against which Harman argues, viz., what he labels 'the foundations theory.'[6] According to Harman, the foundationalist claims that in order to be justified in believing that P...it is required either that P be a foundational belief whose intrinsic justification is not defeated or that there be at least one undefeated justification of P from other beliefs one is justified in believing. If one believes P and it happens that all one's justifications for believing P come to be defeated, one is no longer justified in continuing to believe P, and one should subtract P from one's beliefs.[7]

The foundations theorist is committed to the denial of the principle of conservatism; loosely, this principle states that a belief is justified in the absence of special reasons to think the contrary. According to conservatism, a belief gains defeasible positive epistemic status simply by being believed; every belief is innocent until proven guilty. The foundationalist, on the other hand, asserts that a belief gets no justificatory support from merely being held; a doxastic state is justified only if it is intrinsically justified or soundly based on other justified beliefs. And such justification is, of course, defeasible. This is very important for the foundationalist since she claims that if S justifiably believes P and permissibly infers Q from P, and later comes to have her justification for P defeated, then her justification for Q (as well as any belief justificatorily dependent on Q) is likewise defeated. Thus, new evidence can create a great stir in

one's doxastic pool. Furthermore, the foundations theorist thinks that one ought to always keep track of the justification of one's beliefs. If I believe that P and infer from P that Q, I must continue to keep P in mind if my belief that Q is to remain justified. This point gives rise to the following principle: Principle of Negative Undermining (PNU): One should stop believing P whenever one does not associate one's belief in P with an adequate justification (either intrinsic or extrinsic).

The coherentist, however, rejects the PNU and accepts conservatism. This leads him to replace the PNU with the 'principle of positive undermining.' Principle of Positive Undermining (PPU): One should stop believing P whenever one positively believes that one's reasons for believing that P are no good.

Harman's theory is variety of negative coherentism (despite this, I will frequently follow Harman in calling it simply 'coherence theory'). A belief properly coheres with a doxastic corpus as long as it doesn't conflict with other things that the agent believes; one way in which a conflict can arise is if the agent believes that another belief of hers is based on reasons which are 'no good.' Harman explains that reasons are no good if they are either false or based on false assumptions. Harman's view, then, is that a continued belief is justified as long as it is undefeated. Finally, the coherentist accepts a two-fold goal of belief revision: the maximizing of coherence and the minimizing of change. Although he doesn't couch his theory in these terms, it is clear that Harman takes his theory to be a variety of D-justification. There are two sorts of evidence for this claim. First, the principle of charity requires it. For suppose that Harman's theory were intended as a theory of K-justification. It would then be open to obvious and numerous counterexamples. Here is one: Suppose that (i) upon reflection I find myself believing that my sister's favorite actress is Meryl Streep; (ii) the only reason that I believe this is because I had a conversation with my sister-in-law who told me that Streep was her favorite actress; and (iii) I am now misremembering and thinking that that conversation was with my sister. Surely, my belief that Streep is my sister's favorite actress is not K-justified; should it turn out to be true, I do not know it. However, I do not now have any reason to think that this belief is false, and so the negative

coherentist will have to say that it is justified.[8] Thus, either negative coherentism is not a theory of ongoing K-justification, or it is false. The charitable reading requires us to conclude the former.

The second reason for thinking that D-justification is what Harman has in mind here can be seen in the language of the epistemic principles stated above. The PPU states that one should stop believing a proposition if one sees that its justification is no good. This strongly suggests one has a duty to revise beliefs accordingly, and one has duties to do only that over which one has some degree of control.[9]

After elucidating the coherentist and foundations positions, Harman argues that the coherence theory is superior both as a descriptive and as a normative theory of ongoing justification. He asks us to consider a certain student, Karen, who recently received the results of her aptitude test. Her scores indicate that Karen has considerable aptitude for science and music but little aptitude for history and philosophy. This is surprising to her since she has done well in science and history, and not so well in music and philosophy. In light of the test scores, Karen comes to believe that her history course was easy and that she didn't work hard enough in her music class. Sometime later and much to Karen's chagrin, she finds out that the test scores that she received were not her own and that the results of her tests were lost never to be found.

The first question that Harman asks is descriptive: How will Karen go about revising her doxastic corpus in light of this new information? Harman's answer is surprising. He claims that if Karen is typical, she will continue to hold most of the beliefs that she has recently formed, and reject only that the original report is an indication of her aptitudes; she will continue to believe those propositions that she inferred from her belief about the reports (e.g., that her history class was easy and that she didn't work hard enough in music).

Why does Harman make this counterintuitive claim? In his defense, he cites the psychological literature on a phenomenon known as 'belief perseverance.'[10] The studies of this phenomenon seem to indicate that in very many cases, if a subject comes to believe that P, and on that basis comes to believe that Q, and subsequently has her justification for P thoroughly defeated in a way that she recognizes,

she will nevertheless persist in believing that Q. The application to Karen's case is straightforward: Her initial belief about the report she first received in the mail led her to infer other beliefs about her abilities and from that she inferred beliefs about the difficulty of her courses and the strength of her effort.

So coherentism looks like the better descriptive theory. But surely Karen shouldn't revise her beliefs in this minimalist way; surely, she should take note of the fact that her beliefs about her abilities and her courses depend on the now discredited belief about her test scores.

Foundationalism should be seen as the better normative theory even if it fails as a descriptive account of belief revision. As initially plausible as this claim is, Harman thinks that it must be denied. Let's see why. Harman believes that there is a very simple explanation for the belief perseverance phenomenon: people fail to keep track of the justification of their beliefs. So when S believes P and infers Q, and later has her justification for P undermined, she will not revise her belief that Q because she no longer realizes that the former was her justification for the latter. So Karen cannot be held epistemically accountable for failing to revise her beliefs as the foundationalist would suggest since she doesn't realize that P (the proposition that the report she received reflected her aptitudes) was her justification for Q (the proposition that she is has an aptitude for science and music); therefore, Karen isn't in a position to alter her beliefs about her aptitudes since she doesn't connect P with her aptitude belief. In fact, she might not even continue to believe that P since its significance is exhausted by her belief that Q. All right, one may think, given that Karen hasn't kept track of her justification, we can't expect her to see that there are other beliefs which she should surrender. So there is a sense in which she is justified in believing as she does. However, there is another sense in which her failing to keep track of her justification doesn't let her off the hook. Perhaps she ought to keep track of why she believes as she does. Our pleading ignorance for her now is no excuse; her ignorance is a culpable one. My honestly forgetting that I promised to repay the money that you lent me is no excuse for my not doing what I ought. I should have remembered and my dereliction of this duty surely can't justify my dereliction of the other. Why, then, can't the foundationalist reply to

Harman and the coherentist by saying that Karen is unjustified in continuing to believe as she does because she ought to have kept track of her justifications?

Harman has foreseen such a response and is ready with a reply. He appeals to a principle that the calls 'Clutter Avoidance' (CAP). This principle states that one should not clutter one's mind with trivialities.[11] Writes Harman,

> ...[T]here is a practical reason to avoid too much clutter in one's beliefs. There is a limit to what one can remember, a limit to the number of things one can put into long-term storage, and a limit to what one can retrieve. It is important to save room for important things and not clutter one's mind with a lot of unimportant matters. ...It is...efficient not to try to retain these justifications and the accompanying justifying beliefs. This leaves more room in memory for important matters.[12]

Harman's view, then, is that it is perfectly appropriate and hence justifiable for Karen to lose track of the justification of her beliefs. Humans, after all, are finite beings with limited storage capacity, retrieval ability, and time. Generally, one's capacities are better used if one fails to keep track of the relations between one's beliefs. Harman even suggests that it isn't always good to remember the beliefs upon which another belief is based. Once a conclusion is reached, it is normally appropriate not to remember the premises from which the target belief was derived. Such beliefsare had only as a means to an ends; once they have accomplished their purpose, there is no longer any point in maintaining them.

I believe that Harman is vulnerable to attack on several fronts. I will begin with a pair of objections to his argument for negative coherentism, and finish by arguing that no negative coherence theory is correct as a theory of ongoing justification.

IV. Objections to Harman's Psychological Claims

Harman's argument for coherentism rests on some very implausible psychological claims. It will be the aim of this section to demonstrate that his explanation of the belief perseverance phenomenon and his argument for the CAP are both highly dubious.

a. Belief Perseverance: Recall Harman's explanation of the belief perseverance phenomenon: "..what the debriefing studies show is that people simply do not keep track of the justification relations among their beliefs."[13] The reason Harman thinks these studies show this is that he believes the subjects of the experiments forget that the justification for their new belief about their aptitudes is their belief about what they were told about how they had performed during their ersatz tasks. What makes Harman's claim so very implausible is that in the original experiments performed by Lee Ross and his colleagues, the time between the end of the performance task and debriefing was five minutes.[14] Now in order to have worthwhile results, it is important that the persevering belief be novel. In one experiment, a subject is told that she had done, e.g., well in discriminating genuine from ersatz suicide notes (the alleged performance task), and on that basis she came to believe that she was good at such discrimination. The subject is left alone for five minutes after which the experimenter enters the room and explains that what the subject had been told about her performance on the task had been prearranged and that the experiment was not genuine. Such subjects' beliefs about their abilities to discriminate genuine from ersatz suicide notes tend to persevere in the face of having their original ground discredited. Now Harman's explanation of the belief perseverance phenomenon comes to this: sometime during the five minutes that the subjects are left alone after the task, they forget that the reason they think they are, say, good suicide note detectors is that they have just been told that they have done well on a suicide note detection task. Surely, that is not a credible explanation. However, if they remember the reason for their belief and they now see that it is discredited, then unless the subjects have a new justification for believing what they do, they will be unjustified.[15] b. Clutter Avoidance: My second objection to Harman's psychological claims concerns his use of the CAP.

Immediately following his initial explication of this principle, Harman offers his rationale for accepting it.

To suppose one's mind could become cluttered with beliefs is to suppose such things as (1) that it takes time to add to one's beliefs further propositions that are trivially implied by them, time that might be better spent on other things, and/or (2) that one has 'limited storage capacity' for beliefs, so that there is a limit on the number of things one can believe, and/or (3) that there are limits on 'information retrieval," so the more one believes the more difficult it is to recall relevant beliefs when one needs them.[6]

Let me begin by saying that I think that his first point is absolutely right. If the CAP suggested only that one should not routinely derive everything that one can from one's beliefs, then I would have no qualms with it whatsoever. It is the second and third rationales for the CAP that I find objectionable.

There are at least two ways of understanding Harman's second point; the trouble is that on one way of understanding it, the claim is true but not strong enough to motivate the CAP and the alternative reading has the point sufficiently strong, yet there is no reason to think it is true.

Suppose that we take Harman to be noting simply that humans, being finite entities, cannot store an infinite number of beliefs.[7] Very well, if this is all that is being asserted, I can find little with which to disagree. However, this fact about human memory doesn't support the CAP at all; storage limitation is relevant to the CAP only if there is any real danger of one's long term memory (LTM) being used up. The bare fact that the human brain has storage limitations is in itself no reason to accept the CAP. It is as though one said, "You best not keep your pennies in the Grand Canyon; instead store only your more valuable currency there. For the Grand Canyon is finite in size and so it can hold only a certain amount of money. You should save space for more important denominations." All right, then, suppose we take Harman to be making a stronger, more interesting point, viz., that not only is the human mind finite, but there is a real danger that storage space will eventually be used up. Indeed, the quotation from Harman that introduces the CAP makes it very reasonable to take him to be saying precisely this.

The problem with adopting this reading of Harman's second point is that there isn't any reason to think that it is true. The long term memory store is thought by psychologists to be an enormous receptacle, so large that there isn't any real danger of it ever getting used up.[18]

Harman also argues that the more beliefs one has the more difficult it is to retrieve a significant piece of information when it is needed. He seems to think that the human mind resembles a typical attic. When something needs to be stored, one places it in the attic in the first convenient space. Eventually, the pack-rat will have filled his attic and when he needs to find something important, say the deed to his house, he will have to dig through mounds of junk before he happens upon it. Such a view of memory is totally inadequate. And is very much at odds with the account of LTM that one finds in the psychological literature.[19] There one discovers a picture of memory as that of a library. When one stores a belief in memory one does not simply put it on the first convenient shelf. Rather, one files it with other things to which it is similar; there is a system that determines where the belief will be filed. Brief reflection on human memory is enough to convince one that something like this must be the case. How could we be as successful as we are at accessing relevant memories at a moment's notice if memory were not highly organized? This, of course, does not guarantee that increasing the number of stored items will not decrease recall, but it is sufficient to cast significant doubt on Harman's claim. Finally, it is almost too obvious to need mentioning, but the question here is strictly an empirical one. If we are to accept the claim that adding beliefs to one's doxastic system makes general recall more difficult, we should do it on the basis of solid psychological data. Yet Harman offers no psychological support at all for his claim.

V. Against Negative Coherentism as a Theory of D-Justification

So far, I've offered reasons for thinking that Harman's arguments for coherentism as a theory of ongoing justification leave a great deal to be desired. I shall now offer arguments that show that negative coherentism is a theory that we should not accept.

As I argued earlier, Harman's theory should be understood as an account of D-justification regarding memory beliefs. From this perspective, negative coherentism initially is prima facie plausible. For if I seem to remember that P, and I don't have any good reason to doubt that P, then it would appear that I must be within my epistemic rights in believing as I do. I am not to blame for believing that P even if it turns out that P is false or my belief is poorly grounded. Consider a typical case. Joe seems to remember that his high school calculus teacher's name was Mr. Adams; however, high school was twenty years ago and Joe can't remember why he believes as he does. Nevertheless, he is confident that he is right, and he has never had any significant reason to doubt the reliability of his memory. Now surely it would be wrong to say, with the foundationalist, that because Joe no longer associates his belief with a justification, his belief fails to be justified. On the other hand, if Joe were to come upon strong evidence which suggests that his memory is unreliable in such contexts, or his calculus teacher's name were something different, Joe would no longer be justified in believing as he does. And this is precisely what the negative coherence theory tells us. So what is wrong with it?

A few things. First, in the above case, it was assumed that Joe was not unjustified in coming to believe that his teacher's name was 'Mr. Adams.' But now let's change our assumption. Suppose that Joe never attended his calculus class, and dropped it two days after it began, never to retake it. Joe knew that the only calculus teachers at his high school were Mr. Adams and Ms. Clark, and for no good reason he came to believe that the former was the teacher of his class. Thus, Joe's belief that Mr. Adams was his teacher is initially un(D-)justified. I submit that if this is the history of his present belief that Mr. Adams was his calculus teacher, then Joe's belief is not D-justified. Surely a belief that is initially unjustified doesn't become justified simply because of the passage of time. Yet the negative coherentist is committed to claiming that Joe's initially unjustified belief is now justified (since Joe currently has no defeater for his belief) even though Joe has not acquired any sort of justification in the intervening time.

I realize that he above interpretation of Joe's case is somewhat controversial. It might be true that Joe isn't doing anything wrong when he recalls his belief. That is, if Joe doesn't now remember the genesis of his belief, but only finds himself with the strong conviction that the teacher of the calculus class in which he was enrolled was Mr. Adams, it is very plausible to think that Joe isn't now being epistemically derelict; and so there might be some sense in which it can be said that he is D-justified. Now it is my opinion that if there is such a sense, it is a relatively unimportant one; but I don't intend to argue for that here. For although I believe that this objection to Harman's view is to the point, it is not a particularly devastating one for negative coherentism. Harman can simply amend his account to include a condition requiring the belief to have been previously justified, and to have been undefeated in the interim; indeed, the case of Karen discussed earlier is one in which this additional condition is satisfied.

So is negative coherentism, with the aforementioned epicycle, vindicated? Unfortunately for Harman, the answer is No. We will now examine another, more fundamental problem with negative coherentism (as a theory of ongoing D-justification) that cannot be resolved by a minor adjustment. As I explained earlier, a belief is deontically justified iff one violates no epistemic obligations in holding it. An important issue here concerns the nature of our epistemic obligations. What kinds of acts are we obligated to perform? Harman gives us a clear idea of the type of epistemic duties that he has in mind. We have an obligation to cease believing a proposition if we come to see that our reasons for holding it are no good. We have an obligation not to clutter our minds with unimportant matters (that is, we have an obligation not to believe certain kinds of propositions). And we have no obligation to quit believing a proposition simply because we no longer associate it with an adequate justification. On the other side of the coin, any continued belief for which one no longer has a justification, but for which one also has no reason to doubt, is a belief that one is permitted to retain.

What these have in common is that they are obligations to cease or resist or permission to continue believing a proposition. Since the variety of justification that concerns Harman is deontic, one's failure

to follow these rules will make one epistemically irresponsible and hence blameworthy. However, as mentioned earlier, given a fairly innocuous version of the "'ought' implies 'can'" principle, if one is to be thought blameworthy for failing to comply with such rules, one must be able to comply with them. That is, it must be the case that one's ceasing or continuing to believe a proposition is under one's voluntary control. Hence, Harman's negative coherentism entails doxastic voluntarism, i.e., the thesis that belief is under voluntary control.[20] Yet this raises two questions. To which variety of doxastic voluntarism is the coherentist committed and is the variety of voluntarism in question plausible? The particular voluntaristic thesis that the negative coherentist must embrace is determined by the content of the epistemic duties outlined in his theory. We have seen above that the epistemic principles and rules proposed by Harman are rules which permit or forbid a subject to continue to hold a belief. Do such rules require direct voluntary control over belief, or will indirect control suffice? Let's consider a case. Suppose that Jones believes that P and then comes to realize that he has good reason to think that P is false. Now the normal believer will generally then cease believing that P; but of course that alone doesn't show that she has direct voluntary control over her belief. For to have such control it must be up to her whether or not to continue to believe P. But in a typical case, I think, a person will not be then capable of continuing to believe that P; we are designed in such a way that, usually, if we believe that P and then come to believe that we have evidence which (on the balance) suggests that P is false, we will stop believing as we do. Let's suppose that Jones has indirect control over her belief, and that there are things that she could have done such that, if she had done them, she would not now believe that P; and there are things that Jones can do such that if she does them, she will eventually (but not immediately) no longer believe it. Is a believer in Jones' situation with respect to P one to whom the Principle of Positive Undermining applies? Let's remind ourselves what this negative coherentist principle says:

Principle of Positive Undermining (PPU): One should stop believing P whenever one positively believes that one's reasons for believing that P are no good.

The answer to our question is No. For what the PPU demands is that one stop believing that P whenever one believes that one's reasons are no good. So if there is a time at which a person believes her reasons are no good and yet persists in her belief against her best efforts, then she can't conform to PPU, and hence if the control she has over her belief that P is indirect only, the PPU will state an epistemic principle that will not apply to her.

So the PPU will apply to one only if one has direct voluntary control over one's beliefs; or, perhaps more charitably, PPU will apply to S with respect to her belief that P only if S has direct voluntary control over P. I maintain that this necessary condition of the applicability of PPU makes the principle practically worthless. Although it is not particularly easy to give a argumentative demonstration of the falsity of the direct control version of doxastic voluntarism, none is really needed. Simple introspection can provide ample evidence of the psychological unreality of this thesis. Take any proposition that you now believe, and try, via a simple act of will, to disbelieve it. Of course you can verbally or mentally deny it, you can act as though you believe that it is false, but if you are a human being of standard psychological abilities, you'll find that you can't relinquish it by a simple act of will.[21]

I believe that the falsity of the strong version of doxastic voluntarism does have this important consequence for epistemology: principles that specify conditions of D-justification should not be written as 'belief-rules', i.e., as rules prescribing or proscribing beliefs in particular circumstances. As long as belief is generally not under direct voluntary control, such principles will have little application, at least to human beings. If we can exert only indirect control over our doxastic system, then the epistemic obligations we have can demand only that we do certain things and not that we have certain beliefs. Call rules that tell you what to do (as opposed to what to believe) 'action-rules.' I propose that the right way to think about epistemic responsibility is in terms of compliance with action-rules. This gives us the result that one who, through no fault of her own, believes that P against her best evidence is responsible (i.e., not blameworthy) in believing as she does. One's obligations extend only as far as one's abilities. So if I lack the ability to choose what I believe,

but I am able to gather evidence in a particular way, or to reflect for a given time on certain data, or to remind myself of certain information, or to conduct a certain kind of memory search, or to keep away from the epistemically disreputable, then I can have obligations concerning evidence gathering, memory searches, etc., but not belief production. Keeping this in mind, a belief will be deontically justified as long as I have scrupulously regulated my belief-forming processes. One's believing or failing to believe a proposition will only very rarely, if ever, directly violate (or for that matter comply with) an obligation.

Yet, as the above indicates, this does not mean that deontological considerations have no role in epistemology. The important point here, then, is that while there may very well be epistemic obligations, they will be of no help to the negative coherentist. I can make this point clearer with an example. Suppose that as the result of being raised in a fundamentalist Christian home, Mary believes that drinking alcohol is a sin. Mary is now an adult and while maintaining many elements of her family's religious tradition, she is also rethinking much of what she believes. She decides to look into the biblical evidence for her belief about alcohol. After studying this issue for quite some time, she begins to see that the verses that she has heard her family quote in support of teetotalism have been taken out of context and do not make a good case for what she has been taught. However, despite her recognition of this fact, Mary can't make herself give up her deeply ingrained belief; as they say, old habits die hard. So, for a while at least, Mary continues to believe that drinking alcohol is a sin, even though she also believes that her reasons for holding this belief are not good ones.

The negative coherentist must claim that Mary's belief is unjustified. After all, she is violating the PPU. However, since we are here construing negative coherentism as a theory of D-justification, this simply cannot be the right answer. Mary is surely an epistemically responsible agent vis-a-vis her belief; she has violated no duties. One might even claim that she is epistemically supererogatory. And it is a conceptual truth that if Mary has violated no epistemic duties in forming and sustaining her belief that P, then she D-justified in believing it. Thus, contrary to negative coherentism, Mary's belief is D-justified.

In sum, the root problem with negative coherentism as a theory of ongoing D-justification is that there can be cases of nonculpable belief that fail to cohere (even negatively) with the subject's doxastic system. In order for any theory to be felicitously called 'negative coherentism' it must rule such beliefs unjustified; but if the variety of justification in question is deontic, that verdict is mistaken.

VI. Against Negative Coherentism Simpliciter

At the outset of this paper, I claimed that there are two important and yet significantly different concepts of epistemic justification, viz., that which is closely related to knowledge and that which is fundamentally concerned with epistemic responsibility. While I think that this can hardly be denied, one might be inclined to think that there must be others as well.[22] I will now argue that negative coherentism is generally implausible for reasons that are independent of either of the particular brands of justification we have so far considered. I should make clear, however, that I am not claiming that what follows is anything like a conclusive refutation of negative coherentism. My claim is merely that the considerations to be adduced are sufficiently strong and general to provide us with a rather compelling reason to be suspicious of any variety of negative coherentism as an account of ongoing justification. A primary task of the epistemologist is to develop theories of justified belief. As I argued at the beginning of this essay, such theories will have two parts: accounts of initial and of ongoing justification. While it can be profitable to distinguish these accounts and to begin formulating each independently, it is also necessary to keep in mind that theories of initial and ongoing justification comprise a larger theory, that of justified belief. So one will want to examine each component in light of the other as a check of theoretic unity, or at least, consistency. My claim is that negative coherentism fails to be a good mate for any plausible account of initial justification.

My argument is really rather simple. Either a negative coherentist account of ongoing justification will be paired with a negative coherentist account of initial justification or it won't. Suppose that it is. Then the overall account is theoretically simple, and its parts cohesive. However, the account will have the result that one is

justified in coming to believe a proposition as long as one lacks decisive reasons against it. Now I know of no one who holds such a view, and I think the reason is obvious: it is a lousy theory. It requires neither that one have good evidence, nor that the belief be reliably formed, nor that the subject fulfilled her epistemic duty, nor that the belief is in accordance with the believer's considered epistemic principles. As long as the domain is one about which the subject knows very little (i.e., one with few potential defeaters), she is justified in believing almost anything! It is very hard to see what the point of such a wildly permissive theory of justification would be. Thus, while negative coherentism as a theory of ongoing justification is at home here, the theory with which it is paired is so radically defective that the general account is not worth taking seriously.

Suppose, however, that the negative coherentist regarding ongoing justification prefers a more plausible account of initial justification. So as not to beg any questions, let's say that the theory she adopts is an instance of the following general schema: there is a property f such that S is initially justified in believing that P iff P has f. 'f' could be replaced by 'positively coheres with S's body of beliefs,' or 'is the result of reliable cognitive processes,' or 'is the result of the responsible regulation of S's cognitive resources,' just to name a few possibilities. For whichever account of f we choose, the resulting theory is a conjunction of that with negative coherentism for ongoing justification. Now I maintain that whenever one combines a negative coherence theory with a non-negative coherence view of initial justification, one will get an uneasy alliance; or to change the metaphor back to the one I've been working with, the parts of the whole will not fit squarely together. Two problem cases demonstrate this. First, suppose that at t S forms a belief that P, but that P fails to instantiate f. So S's belief is unjustified; but once it is formed, negative coherentism as a theory of ongoing justification will imply that as long as S is aware of no good reason to think P is false or unreliably formed, her belief is justified. So having formed an unjustified belief, her belief will become justified simply by being believed. We have already looked at an example of this kind of case in the discussion of Joe and his belief about his calculus teacher. There I suggested that Harman could avoid this problem by tacking on as a

condition of ongoing justification that the belief must have previously become justified and been undefeated in the interim. While this does alter the view significantly, it shares enough of the original's features to be plausibly construed as the same basic theory. Even if we amend the position accordingly and thus diffuse the above case, there is still another problem. Suppose that S forms her belief that P at t, and that at t, P has f, and is thus justified. However, a moment later, once the belief is formed, it no longer has this property. So suppose, e.g., one is an evidentialist, and one claims that one is justified in forming a belief iff one has good evidence for it. And suppose that S comes to believe that P on the basis of a great deal of evidence and is justified in her belief, but that she soon forgets all of the evidence she once possessed. We can even suppose that her amnesia is such that she also is no longer aware of any evidence that her memory is generally reliable. In short, she believes that P, but hasn't a shred of evidence for it. To claim that her continued belief is justified is completely foreign to the spirit of evidentialism. Why would the evidentialist have such a wildly different view of ongoing justification? Or suppose that the theorist is a reliabilist, and that the belief in question is formed on the basis of a reliable cognitive process. However, suppose furthermore, that the person's memory is extremely unreliable, and that a short time after forming the belief, she recalls it. Now the reliabilist will surely not want to say that her belief is justified if undefeated; for the cognitive relation to truth that is near and dear to the reliabilist's heart has been severed if memory is unreliable. So it seems that the negative coherence theorist of ongoing justification faces the following dilemma: she must either adopt an account of initial justification that stands little chance of appealing to anyone, or adopt an account of initial justification that is theoretically at odds with her favored theory of ongoing justification.

Again, I do not assert that no such theory of ongoing justification can get around this dilemma. To make good on that claim, I'd have to show that there is no concept of justification which can be analyzed along negative coherentist lines. I have not done that. However, I am unaware of any variety of justification for which negative coherentism is plausible as an account of initial justification; and I'm unaware of any theory of initial justification (which has adherents)

the spirit of which is compatible with a negative coherentist account of ongoing justification.

VI. Conclusion

Negative coherentism has a certain undeniable prima facie attractiveness as a theory of ongoing justification. However, I believe that our discussion has shown that there is no clear place for such a theory. It rather obviously fails to capture our intuitions about K-justification, and upon examination, it is also proves inadequate as a theory of D-justified belief. Furthermore, there is ample reason to suspect that it is inadequate as a theory of any variety of ongoing justification. [21]

THE UNIVERSITY OF ARKANSAS
FAYETTEVILLE, ARKANSAS 72701
U.S.A

NOTES

[1] In a previous article, I have examined internalistic foundationalist theories of ongoing justification. See Thomas D. Senor, "Internalistic Foundationalism and the Justification of Memory Belief," *Synthese* 1993.

[2] Gilbert Harman, *Change in View* (Cambridge, MA, 1986), Chapter four.

[3] I should note that Harman never uses the expression 'ongoing justification' to denote what he is discussing; rather, he calls his view a theory of belief revision. Despite this terminological difference, I believe that the issue he is concerned with and that which is my focus are essentially the same.

[4] For the purposes of this essay, I am taking 'duty' and 'obligation' to be synonymous. Also, when I use these terms I have in mind what is sometimes called 'subjective duty.' That is, what is essential to the concept of deontological justification is epistemic blamelessness; if one's view is that objective duty and praiseworthiness/ blameworthiness can come apart, it is the latter that concerns the notion involved. Failure to recognize this has led Paul Moser and Matthias Steup to offer unconvincing defenses of the deontological conception against attacks by William P. Alston (see note 5). See Moser's *Knowledge and Evidence* (Cambridge University Press: New York, 1989) pp. 38-40 and Steup's

"The Deontic Conception of Epistemic Justification," *Philosophical Studies* 53, no. 1 (1988).

5 See William P. Alston, "Concepts of Epistemic Justification" *The Monist* 1985, pp. 57-89.

6 It should be noted at the outset that Harman recognizes that he is using the terms 'foundations theory' and 'coherence theory' in nonstandard ways. The primary deviation from standard usage is that the theories Harman is discussing are theories of belief revision or what I am calling 'ongoing justification.' They don't pretend to be more general theories of justification or positive epistemic status.

7 Op. cit., page 31. 8
If this case isn't convincing, I invite the reader to make up one of his own. It's easy. The sort of justification that does most of the work in converting true belief to knowledge must be strongly truth connected. However, there is no such restriction on a negative coherentist's view of justification. So it is clear that one can have an occurrent memory belief that P, have no reason to believe that P is false, and yet not believe P in a truth-connected way. In such an event, one will have a belief that is justified according to Harmanian negative coherentism but which is not justified in the sense needed for knowledge. Of course, such belief might be epistemically blameless; whether this is so will be considered presently.

9 In his paper "Epistemic Obligations," (in James Tomberlin, ed., *Philosophical Perspectives 2: Epistemology* 1988 (Atascadero, CA: 1988)), Richard Feldman argues that there are obligations that do not require the ability to satisfy them. One has to agree that there are such obligations (i.e., financial and academic), but it is far from clear that epistemic obligations are among them.

10 A good survey of the belief perseverance phenomenon (and the one Harman quotes from) is "Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments" by Lee Ross and Craig A. Anderson, in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge, 1982), pages 129-152.

11 Ibid, page 12.

12 Ibid, pages 41 and 42.

13 Ibid, page 38.

14 See L. Ross, MR. Lepper, and M. Hubbard, "Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the

Debriefing Paradigm", *Journal of Personality and Social Psychology*, (1975), pp. 880-892.

[15] It might well be that they do now have a new justification. A more plausible explanation of what is going on here is that during the five minute period, subjects scan their memories for confirmation of their new belief. They will likely find what they take to be confirming cases, and it is likely that no disconfirming cases will come to mind (this tendency is known as the 'confirmation bias'). So by the time they are debriefed, they have before their minds what they take to be independent evidence of the truth of their new belief, so when the old justification is discredited, the new one still remains.

[16] Ibid, page 12.

[17] Here one needs to make the distinction between explicit and implicit beliefs. An explicit belief is a belief which "involves an explicit mental representation whose content is the content of that belief," (Harman, page 13) while a belief is implicit if it is not explicit but is "easily inferable" from others that are. So what is being claimed is that since one has limited storage capacity, one can store only so many explicit beliefs.

[18] For example, Michael W. Eysenck and Mark T. Keane, in *Cognitive Psychology: A Student's Handbook* (Lawrence Earlbam Associates: UK, 1990), write: "Despite the complexities associated with assessing the capacity of the short-term store, there is general agreement that its capacity is strictly limited. This is a major difference between it and the long-term store, for which there are no known limits" (p. 139). Further evidence for the claim that psychologists do not think of the long-term memory store as limited in the way Harman claims can also be found in the Eysenck and Keane text: In spite of these uncertainties, many prominent psychologists from Freud to Tulving have favoured the hypothesis of permanent storage. This hypothesis can be expressed in the following fashion [(Loftus, E.F. & Loftus, G.R., "On the Permanence of Stored Information in the Human Brain," *American Psychologist*, 35, 409-420)]: "Everything we learn is permanently stored in the mind, although sometimes particular details are not accessible." ..Loftus and Loftus [(op. cit., p. 410)] conducted an opinion poll among psychologists, and discovered that 84% of them endorsed this statement. I take it that if psychologists take the permanent storage thesis this seriously, and yet refrain from offering Harman's advice, they must presume that there are no practical limitations to what can be put in LTM. For other discussions that clearly reject the thesis that there is any

practical limitations on LTM, see John R. Anderson, *Cognitive Psychology and Its Implications* (Second Edition) (W.H. Freeman: New York, 1985), Chapter 6, and Alan Baddeley, *The Psychology of Human Memory* (Basic Books: New York, 1976), Chapter 6.

[19] The psychologists quoted in the previous footnote, Eysenck and Keane, also claim that "Those who learn most efficiently are generally those who already have the greatest amount of information stored in long-term memory. This is exactly the opposite of what would be expected if the long-term memory store had a relatively modest capacity" (op. cit., p. 139). Since learning is determined by successful recall performance in experiments, Harman's assertion that recall ability is retarded by an increase in stored information turns out to be precisely contrary to what is the case.

[20] For the best discussion of doxastic voluntarism of which I am aware, see William P. Alston, "The Deontological Conception of Epistemic Justification" in Tomberlin, op. cit., pp. 257-299.

[21] I can think of precisely one general variety of simple acts of will within in my control which are such that if I perform them, I'll immediately cease to believe some proposition P. An example of this act-type is the simple act of pulling the trigger of a loaded gun whose barrel is placed at my temple; in that case, one might suppose, I should immediately cease believing all propositions.

[22] For example, the concept that occupies Richard Foley in *The Theory of Epistemic Rationality* (Cambridge, MA: Harvard University Press, 1987) would seem not to be one of the notions I've glossed. Also, the technical literature on belief revision (e.g., Peter Gardenfors' book *Knowledge in Flux: Modeling the Dynamics of Epistemic States* (Cambridge, MA: MIT Press, 1988) would appears to be about something other than K- or D-justification.

[23] Thanks to Rudy L. Garns, Alvin Goldman, Asa Kasher, Mark Klinger, Del Ratzsch, Scott Sturgeon, and Nicholas Wolterstorff for helpful comments on earlier versions of this paper.