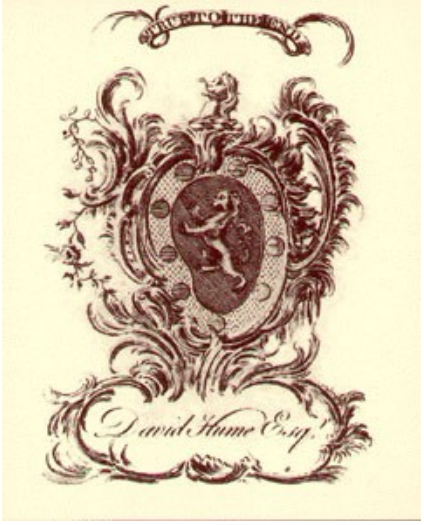

Stability and Justification in Hume's Treatise, Another Look— A Response to Erin Kelly, Frederick Schmitt, and Michael Williams

Louis E. Loeb

Hume Studies Volume 30, Number 2, November (2004) 339 - 404.

Hume Studies
Volume 30, Number 2, November 2004, pp. 339–404

A Symposium on Louis E. Loeb,
Stability and Justification in Hume's Treatise

Stability and Justification in Hume's *Treatise*, Another Look— A Response to Erin Kelly, Frederick Schmitt, and Michael Williams

LOUIS E. LOEB

The symposiasts press from a number of directions.[1] Erin Kelly contends that Hume's stability-based sentimentalist ethics cannot do justice to our considered normative moral judgements. Schmitt and Williams criticize my account of Hume's epistemology proper. I will have to give ground: my book does overstate the extent to which Hume reaches a destructive result, in large part because I overlook significant variants of a stability account of justification. I make other concessions—in regard to the country gentlemen passage and Hume's 1.3.9 treatment of resemblance—but believe these have limited repercussions.

Let me take note of some large-scale features of the debate with Schmitt and Williams about Hume's theory of justification. We share a number of fundamental theses:

- Hume has a sustained, constructive project of drawing normative epistemological distinctions in terms of a purely "naturalistic" theory of justification.
- The constructive project is well in place in 1.3. There is a corollary: 1.3 is not confined to cognitive psychology.

Louis E. Loeb is Professor of Philosophy, University of Michigan, 435 South State Street, Ann Arbor, MI 48109-1003, USA.
e-mail: lloeb@umich.edu

- Causal inference enjoys epistemic pride of place within the constructive project. Another corollary: in part 3, there is an argument against "Reason" as traditionally conceived, but no "problem of induction" beyond that.
- Hume's basis for his epistemological distinctions cannot be derived from such notions as irresistibility or involuntariness; the resources of the Kemp Smith tradition of interpretation are too thin to account for Hume's normative epistemology,
- Any skepticism in the *Treatise* (apart from the attack on Reason) arises within the naturalistic epistemology itself.

Some of these shared claims go against the grain of major lines of interpretation. At the same time, there are deep disagreements: about whether there is so much as a destructive turn in 1.4.7 and about the character of Hume's constructive epistemological position. Schmitt urges that the stability reading cannot account for the veritistic elements in Hume's texts, suggesting that they better fit a reliabilist interpretation. Williams is on board with a stability interpretation, but insists that Hume's concern is long-term consensus rather than stability for the individual.

This response is divided as follows:

I. Reply to Kelly

1. Stability, Reflective Endorsement, and the Motivation for the Steady Point of View
2. The Narrow Circle, Invisibles, and Partiality
3. Information, Sympathetic Templates, and Stigmatized Groups
4. Common Morality and Roles for Reflection

II. Reply to Schmitt

1. The Stability Account of Belief and the Aim of Truth
*2. Actual versus Reflective Stability
*3. Peak Stability versus Average or Temporal Stability
4. Metaphysical Beliefs and the Identity Propensity

*III. Hume's Mix of Stability and Truth

IV. Reply to Williams

*1. The Two-stage Model
2. "External World Scepticism"
*3. *Treatise* 1.4.7
*4. Long-term Consensus

Part 3 of my reply to Schmitt incorporates his fruitful options for a stability interpretation and thus repairs shortcomings in my reading. The resulting revisions are pivotal to parts 1 and 4 of my reply to Williams. Part 3 takes up broad matters of interpretation germane to both Schmitt and Williams.[2] The asterisked material

constitutes an abbreviated route through the discussion of Hume's epistemology and metaphysics.

I. Reply to Kelly

1. Stability, Reflective Endorsement, and the Motivation for the Steady Point of View

Kelly reminds us that sentimentalist moral theories are at risk of licensing moral judgments that we would reject as resulting from "distortions," failures to give all persons "full sympathetic attention" (Kelly 329, 335). First, there are "outsiders" (Kelly 334, 335); we care more for persons inside our group or groups (Kelly 329), generating garden variety cases of partiality. Second, there are outsiders whom we regard with "a sense of their unfamiliarity" or "difference," lowering or degrading their status as objects of sympathetic response (Kelly 333, 334). Such stigmatization—from "prejudice," "propaganda that aims to dehumanize," and "racial stereotypes" (Kelly 334)—goes beyond mere partiality. Not all outsiders are stigmatized; foreign nationals, for example, need not be associated with differences which carry a negative valence.[3] Finally, there are outsiders whom we effectively ignore; they are shielded or blocked from the reach of our sentiments, "beyond our sphere of awareness" (Kelly 335; cf. 333–4), leaving them, as it were, without any moral standing at all. In sum, we can distinguish mere outsiders, the stigmatized, and invisibles.

Kelly compares the merits of the stability and reflective endorsement interpretations of Hume's theory of the normativity of moral judgment. Kelly "offer[s] no decisive verdict" (Kelly 335), though she seems prepared to give the stability interpretation the nod (cf. Kelly 335–6). Kelly argues that there is a price to pay: the problem of distortions "is especially urgent" insofar as Hume conjoins sentimentalism with the stability account of normativity (Kelly 329–30); a sentimentalism embedded in a reflective endorsement account might better avoid the problem. The generous balance in Kelly's discussion should not obscure the difficult challenges she raises for Hume's sentimentalism and for the viability of the stability interpretation.

Both the stability and reflective endorsement interpretations apply to Hume's account of the normativity of the understanding and morality alike.[4] It is Korsgaard's own view that the understanding fails the test of reflective endorsement in 1.4.7, so that her reflective endorsement account would seem to undermine the understanding tout court. She writes: "Skepticism about the understanding is in order. . . . Hume clearly thinks that the understanding fails this [reflexivity] test."[5] Korsgaard undertakes to contain this result, I think unsuccessfully (SJ 90n8). By contrast, the stability interpretation more easily accommodates the positive normative status of the understanding in the case of beliefs that are held without intense reflection (cf. Kelly 336 and part 3 of my reply to Schmitt). To the extent that it is implausible to interpret Hume as identifying the normativity of the

understanding with reflective endorsement, Korsgaard's interpretation of the normativity of the moral sense loses the luster that derives from generality.

If we confine our attention to the moral sphere, Korsgaard is on stronger ground: the test of reflective endorsement issues an unambiguously positive result in the conclusion of Book 3, if not in the conclusion of Book 1. But I think Kelly is correct that the text does not support Korsgaard's interpretation (Kelly 336). The sense of morals, Hume writes, must "acquire *new* force, when reflecting on itself, it approves of those principles, from when it is deriv'd" (T 3.3.6.3, emphasis added; SBN 619). This passage is Korsgaard's only direct evidence (evidence other than the generalization from the case of the understanding) for locating in the *Treatise* the reflective endorsement account of morality.[6] The natural reading of 3.3.6 is that reflection generates *additional* normative standing, so that there must be a source of normativity other than reflective endorsement.

As Kelly nicely emphasizes, on the stability interpretation no reflective endorsement test is required for either positive or negative normative standing (cf. Kelly 331–3). The stability interpretation explains the attraction of the reflective approval interpretation, without endorsing it (SJ 26–8, 97). Reflective endorsement is a special case of stability, failure in reflective endorsement a special case of instability; the normative standing of reflection is derivative from any stability or instability it fosters. A reflective endorsement interpretation inverts the relationship of stability and reflective endorsement. Kelly presses the point that even if normativity does not consist in reflective endorsement, the stability interpretation leaves open the possibility that Hume allows that there are kinds of reflection that enhance normative standing by contributing to stability (Kelly 333, 334). This is exactly what I want to say.

Although I disagree with Korsgaard's interpretation of Hume's account of normativity, we agree that we adopt the "*steady* and *general*" point of view in order to avoid "*contradictions*" (T 3.3.1.15; SBN 581, cf. 3.3.1.18, 3.3.3.2; 583, 602) in our judgments. The nature and role of these "contradictions" constitutes a central interpretive issue (SJ 119–22).[7] Suppose we construe the variation in judgment due to the operation of sympathy as merely reflecting variation in the apparent moral qualities of actions or characters from different perspectives, points of view that depend upon different relationships to the agent under evaluation and those affected by him. The "contradictions" vanish if variation in judgment simply reflects differences in seeming or apparent virtue or vice. Why not invoke this perspectival model?

My book provides an interpretation of Hume's answer (SJ 122–3, 127–8). Although the judgments that directly reflect variation in sentiment do not constitute strict contradictions, they *feel like* contradictions; they produce a feeling of uneasiness or dissonance of much the same sort produced by genuine contradictions. In virtue of a general aversion to pain, this dissonance supplies a motive to remove

or mitigate the associated discomfort. Korsgaard stresses that differing judgments generate "a sense of contradiction," "commotion," "instability," "tension," "conflict" that "must be quieted," and thus "pressure" toward adopting a steady point of view.[8] Our interpretations diverge in matters of detail (SJ 128n40), but we agree that Hume takes the discomfort arising from a feeling of contradiction to exert pressure toward a point of view in which judgment is stabilized.

We differ in the normative significance of this psychological explanation of the adoption of the moral point of view. For Korsgaard, absent reflective endorsement, normativity has not set in. On the stability account of normativity, the instability of our more spontaneous, "contradictory" judgments of virtue and vice detracts from their normative status; positive normative status accrues to the stable judgments in the steady and general point of view and does so even though reflection need not have played a role. This explains how Hume can think reflective endorsement adds to normative standing already in place.

We need not resolve the issues between stability and reflective approval interpretations in order to pursue a question Kelly suggests: in Hume's account, to what extent does reflection enhance the stability of moral judgment—and therefore *at least contribute* to the normative status of those judgments? Beyond the general reflection on the source of moral distinctions at *Treatise* 3.3.6, does reflection play any essential role in stabilizing moral judgment? For example, might reflection contribute to Hume's solution to the problem cases Kelly advances (cf. Kelly 333–4)? At first blush, the account of the psychological mechanism that leads to the adoption of a steady and general point of view assigns no essential role to reflection; the mechanism consists in the pressure of discomfort or uneasiness that arises unreflectively. A negative answer to our question would nevertheless be premature. True, reflection plays no essential role in *motivating* the adoption of a steady point of view, but we have not explained how we *succeed in adopting* it. So there yet may be a role for reflection. I return to this question in part 4, after considering the content of the steady point of view and its bearing on Kelly's problem cases.

2. The Narrow Circle, Invisibles, and Partiality

The account of why we adopt a steady and general point of view is as far as I take matters in my book. I do not there identity the point of view that respects the stability in moral judgment. Is it that of an observer who sympathizes with other persons as if he stood in the same relationship to each one; or is it the point of view of someone who stands in a specified relation to the agent; or what? Until such questions are addressed, it is not possible to spell out the normative implications of Hume's sentimentalism.

Kelly outlines an answer: the steady point of view is that of an observer who sympathizes with the benefits and harms that a character trait tends to produce

within the agent's narrow circle (Kelly 331). There are two elements here: the moral judge considers the effects of a character trait specifically on the agent's narrow circle; and the judge relies on "general rules" about a character trait's *tendency* to produce benefits and harms. This latter element, the burden of Hume's discussion of "virtue in rags" (T 3.3.1.19; SBN 584), is uncontroversial.

There is also no question that Hume takes the agent's narrow circle to be implicated in the steady point of view: "we cannot afterwards fix ourselves so commodiously by any means as by a sympathy with those, who have any commerce with the person we consider" (T 3.3.1.18; SBN 583). Psychological salience explains why a focus on the agent's narrow circle has pride of place among various fixed sympathetic perspectives. How is this focus to be understood? According to Kelly, moral approval is determined by the sentiment "a 'judicious spectator' . . . would feel when considering the tendency of [character] traits to benefit . . . persons who have a 'particular connection' with the agent" (Kelly 331). I take Kelly to mean that the judicious spectator *sympathizes with* the typical effects of the character trait on the members of the agent's narrow circle.[9] Since Hume makes a number of statements exactly along these lines (T 3.3.1.18,28,30, 3.3.3.2; SBN 583, 590, 591 602–3), this first element in Kelly's sketch is often attributed to Hume.

This reading cannot be correct. It does not specify the relationship between the judicious spectator and the agent's narrow circle, so that it leaves variability in sympathy in place. An observer might stand in any of a number of relationships to members of that circle: friend, acquaintance, fellow countryman, or that of an "indifferent" (T 3.3.2.8; SBN 488) person. To take cases prominent in Kelly's discussion, the agent's narrow circle might consist of persons stigmatized for a spectator, or invisible to a spectator, and so forth.

There is another way to construe the role of the narrow circle. Suppose Hume's thought is that the judicious spectator adopts a point of view in which the pleasures and pains that register on him most strongly are those of the agent's narrow circle. We can then exploit the doctrine that sympathetic response varies with special resemblances and relations (cf. T 2.11.5; SBN 318)—distance in psychological space. Were a moral judge to adopt a point of view far removed from that of the agent's narrow circle, the pleasures and pains of the narrow circle would not register strongly. In order for the pleasures and pains that register most strongly to be those of the agent's narrow circle, a moral judge must sympathize with its members from close-up, either from nearby or from within the narrow circle itself. Adopting the latter option, the judicious spectator sympathizes most intensely with the agent's narrow circle in virtue of sympathizing from within that circle. The steady and general point of view is that of an observer who sympathizes, from within the agent's narrow circle, with the pains and pleasures that are the usual effects of the agent's character. This is the sympathetic perspective that insulates moral judgment from variation in sympathetic response.[10]

Hume's account of the steady point of view avoids Kelly's concern that the character traits of invisibles "do not constitute virtues or vices," since "we do not respond sentimentally to persons beyond our sphere of awareness" (Kelly 335). Consider a prisoner, perhaps secretly held. Much as the person in rags, the prisoner is in an unusual set of circumstances. This does not deprive him of a narrow circle, consisting of some mix of fellow prisoners and (even if he is in isolation) of family and friends.[11] Virtue and vice is a matter of the sentiment that would be produced from the point of view of a member of that circle. The prisoner is not in any relevant sense invisible or outside the sphere of awareness of his narrow circle. On Hume's theory, the prisoner does have virtues or vices, as the case may be. Those of us well removed from the prisoner's narrow circle might not know what the sympathetic response of a member of the prisoner's circle would be, and hence might not know whether the prisoner is virtuous or vicious. This is as it should be, since we lack relevant beliefs about his traits and dispositions. (This is not yet to address the question of our moral obligations *to* invisibles.) Of course, as discussed in the next part, even members of the agent's narrow circle may lack information relevant to sustain a moral judgment.

What about Kelly's cases of partiality in our judgments about obligations to others? Moral judgment reflects the sentiment we would experience were we to sympathize, from the perspective of the narrow circle of the agent, with the usual effects of his character. A character trait might tend to have effects—usual or characteristic effects—upon persons far removed from the agent, such as strangers or foreigners. Sympathy, which operates in virtue of the "great resemblance among all human creatures" (T 2.1.11.5; SBN 318), does have this universal reach. The pleasures and pains of everyone (everyone typically affected by a character trait) thus carry some weight in the moral point of view. (No one need be invisible to the moral judge in virtue of the construction of the steady point of view.)

On the other hand, sympathetic response is heightened by special relationships: "any peculiar similarity . . . facilitates the sympathy" (T 2.1.11.5; SBN 318). Sympathetic response depends upon the proximity or closeness in psychological space between the steady and general point of view, that of the narrow circle of the agent under evaluation, and the person with whom an occupant of that point of view is sympathizing. It is here that Hume's psychological doctrines shape the normative implications of his sentimentalism. The moral judge sympathizes most intensely with other members of the agent's narrow circle, less intensely with those at one remove from that circle, and so forth. The pains and pleasures of a fellow member of the narrow circle receive more weight than those of countrymen outside the narrow circle, who themselves receive more weight than those of someone from another country. The theory is universalist in that the pains and pleasures of every relevant person (every person typically affected by the agent's character) count in the sentimentalist calculus. Persons, however, do not count equally; the

moral relevance of their pains and pleasures is proportional to their psychological proximity to the *agent's* family and friends. The result is a weighted, universalist moral theory—weighted toward the agent's narrow circle. Considerations of stability, in conjunction with the doctrines about the operation of sympathy, have partiality for a consequence.[12]

This result bears on worries about the implications of Hume's sentimentalism with respect to moral obligations to outsiders, persons removed from the agent's narrow circle. Hume's weighted universalism is unabashedly partialist in the sense I have identified. Hume sees this sort of partiality as characteristic of moral judgment (cf. T 3.2.1.18, 3.2.2.8, 3.3.3.2; SBN 483–4, 488–9, 602–3). Partiality in distinctively moral judgment is, in Hume's view, part of the data for which a theory of moral judgment ought to account. For this reason, he could claim the implications of his position—partiality to those within the agent's own narrow circle—as confirmation of a sentimentalism embedded within a framework that locates normativity in stability.

Kelly writes of "full sympathetic attention" (Kelly 335), a sentimentalist version of a substantive requirement of "due consideration for all persons" (Kelly 337). If "full" or "due" consideration means equal consideration, this requirement is not Hume's. As Kelly recognizes, "some sentimentalists reject a conception of morality as impartial and universal" (Kelly 329). Once this requirement is rejected, the degree to which morality requires impartiality is an open question. Hume's sentimentalism, together with his account of normativity, supply a principled answer: because the perspective from within the agent's narrow circle is salient and hence stable, it constitutes the steady and general point of view; and this point of view yields partiality in moral judgment, owing to the intensification of sympathy by special relationships.

Let me offer two comments. First, although the moral point of view is partial to the agent's narrow circle, Hume's account provides a clear sense in which moral judgment is "impartial"; the moral judge adopts the sympathetic perspective of the agent's narrow circle, and does so whether or not he is himself a member of that circle or bears any special relationship to its members. In the moral point of view, we do transcend our own concerns and interests, except insofar as we happen to bear a special relationship to the agent and his narrow circle. Moral judgment is impartial in that, *qua* moral evaluator, one's interests are left out of account; if they come into account, it is only because the evaluator happens to stand in a special relationship to the *agent's* narrow circle.

Second, partiality in ethics has, in recent decades, gained approval within a variety of theoretical perspectives—virtue, communitarian, and feminist ethics among them. Hume's development of sentimentalism within a stability account of normativity offers an interesting, and by my lights appealing, foundation for a partialist morality. Kelly, however, would contend that any attractions of Hume's

position are, at the least, marred by difficulties in doing justice to our moral obligations to persons who are invisible or stigmatized. I tackle these matters in the next part.

3. Information, Sympathetic Templates, and Stigmatized Groups

"[B]eing nothing but a lively idea converted into an impression" (T 2.2.9.13; SBN 385–6, cf. 2.1.11.4,8; 317, 319), sympathy cannot operate in the absence of relevant beliefs. These may or may not be justified and may or may not (if justified) constitute knowledge. Using cognates of 'information' broadly to refer to any belief: to what extent must the judicious spectator be informed—about the agent's traits and dispositions and about their usual or characteristic effects (and hence about sociology and psychology)?[13]

Any answer bears on Kelly's concerns about the normative implications of Hume's theory in regard to moral obligations to persons who are invisible or stigmatized (as distinct from moral evaluation *of* invisibles, discussed earlier). Kelly has in view socially generated "tendencies" to block out or to distort sympathetic response. These tendencies result from such circumstances as physical removal and official silence in the case of invisibles, and propaganda and stereotyping in the case of stigmatized groups (cf. Kelly 333–4). To the extent that blocking and distortion result in defects in information, a better informed spectator will be less susceptible to "distorted" responses.[14] There is a substantial body of psychological literature that takes prejudice to be responsive to information, even if it is rooted in misinformation. Hume does not characterize the information incorporated into the steady and general point of view, information that could be incompatible with the ignorance and misinformation in Kelly's examples. Absent such a specification, the deliverances of Hume's normative theory, and hence the force of Kelly's counter-examples, are indeterminate.

We can try to bring matters into some focus. The broad question is how much information is possessed within the steady point of view, possessed by a moral judge within the agent's narrow circle. Questions of this form are familiar within the context of *ideal* observer theories, where it is natural to equip the observer with full or complete information. Were Hume to set this demanding requirement, the problem of invisibles would be solved, and the problem of the stigmatized would dissipate, to the extent that prejudice recedes under the pressure of information. (I take up cases of bias impervious to information below.) A requirement of full information, however, is not Humean in spirit. In the course of considering the obligation of benevolence for a moral agent, Hume writes: "Being thus acquainted with the nature of man, we expect not any impossibilities from him" (T 3.3.3.2; SBN 602). Presumably, Hume would also not expect epistemic impossibilities from the moral judge.

We might suppose that Hume equips the moral judge with the beliefs that a member of the agent's narrow circle would normally possess; the judge's information would closely match that of a typical member of the narrow circle, where ignorance and misinformation could be widespread. Were Hume to set this modest requirement on information, his normative theory would be susceptible to many of the distortions in Kelly's examples. We are thus presented with a range of options; the informational requirement can be ratcheted up (as with an ideal spectator), tamped down (as with the modest requirement), or set somewhere in between.

The modest informational requirement is problematic independently of Kelly's cases of invisible and stigmatized groups. The point of Hume's construction is to identify a point of view that will best mitigate, if not remove, variability in judgments about character. We are similarly at risk to fluctuation in information, especially when information we do not possess is readily available to us. If moral judgment were tied to the beliefs that a member of an agent's narrow circle would typically possess, it would be vulnerable to conflict with judgements that arise in the context of beliefs that could easily be acquired (cf. part 2 of my reply to Schmitt). Judgments based on typically possessed beliefs would be prone to conflict with the judgements of others based on readily available information, as well as to conflict with one's own future judgments as one accesses such information. It seems odd to expect a moral judge, in adopting the steady and general point of view, to suppress information that could be prevalent within the agent's narrow circle itself. For these reasons, the modest informational requirement is not the most salient response to the problem of mitigating variability due to differences in information. The pressures toward stability motivate incorporating readily available information into the construction of the steady point of view.[15] The *judicious* spectator has less information than an ideal observer, but more information than is typical of a member of the agent's narrow circle. This informational requirement has the potential significantly to remedy the effects of ignorance and misinformation in Kelly's examples.

Kelly emphasizes that the conditions responsible for lack of information and misinformation impose "fairly stable limits on the range of social circumstances in which we encounter difference and disagreement" (Kelly 334); the mechanisms in play "are stable enough to block our sympathetic identification with" (Kelly 334) persons who are invisible or stigmatized. These comments are directed at a misleading conception of the way in which uneasiness motivates the adoption of the steady and general point of view. On Kelly's picture, there is no motivation to adopt the steady point of view in individual cases of untoward normative judgments that are *de facto* quite stable. For Hume, however, variability in point of view, much as variability in judgment, is unsettling. Adoption of the steady point of view is thus itself a "general rule" (T 3.3.1.20–21; SBN 585, cf. 3.3.3.2; 602–3), even where there is little or no variability (or even little likelihood of variability) in judgment due to differences in information. Adopting the point of view of the moral judge is not a

case by case expedient. Stable social circumstances may conspire in the direction of invariability in response with respect to particular normative judgments; it remains that the judicious spectator relies on readily available information.

But wait. What if sympathetic response to some groups is impeded by prejudice that is not supported or sustained by misinformation? Suppose, for example, that bias against persons with a purely cosmetic disfigurement arises or persists even absent misinformation. In Humean terms, we can suppose that the physical dissimilarity degrades sympathetic response and thus distorts moral judgment. It is of the nature of Hume's weighted universalism to blunt the effects of such prejudice. Consider the virtues of benevolence and generosity. To the extent that persons with the disfigurement are members of the agent's narrow circle, the ties of family and friendship can be expected to assert themselves and at least predominate.[16] To the extent that they are strangers, outside an agent's narrow circle, bias toward them will effect what is already a small pie, since virtue requires relatively little generosity to strangers (part 2).

Hume has another line of response: in generating basic moral judgments, the judicious spectator sympathizes through a particular set of screens, abstracting from many facts about individuals that trigger prejudiced responses. Hume's persistent emphasis on the agent's relations to particular groups—friends, acquaintance, countrymen, and foreigners (T 3.2.2.6,8,11, 3.2.7.1, 3.3.1.14,17,18; SBN 487, 488, 491, 534, 580–1, 582, 583, cf. 1.3.9.13; 114)—suggests that the moral judge sympathizes with others in a particular way. Evidently, in sympathizing with every person in the universe who can be expected to be affected by an agent's character trait, the judicious spectator does not take into account all the details of each person's special relationships to the agent's narrow circle; rather, he sympathizes with individuals *qua* relative, friend, acquaintance, etcetera. This moves in the direction of the view that the moral judge sympathizes by way of prototypes or stereotypical representatives of these groups. We also have special relationships to persons with the same surname or of the same trade (T 2.2.6.2; SBN 352). Yet, in his more theoretical discussions of moral judgment, Hume does not highlight sympathy with the members of these groups. Similarly, he does not attend to the myriad of other communities defined on the basis of geography, religion, profession, and the like. These examples highlight the need to motivate Hume's focusing on some groups rather than others.

Although Hume does not explicitly address this question, we can offer an hypothesis to explain his specification of the groups that are material within the moral point of view. Readers of the *Treatise* are unlikely to balk when Hume highlights family, friends, acquaintance, fellow countrymen, foreigners.[17] These groups are salient in virtue of the convergence of two features: first, they are nested, ordered by set-inclusion (or close enough); second, this nesting coincides with a psychological ordering, in that we sympathize most intensely with the least

inclusive groups.[18] Fellow tradesmen overlap the other groups; if we add persons of the same trade to the list of categories, the double nesting is lost. We similarly sacrifice the double nesting if we add members of the same community—our local jurisdictions, coreligionists, coworkers, and so forth.[19] It is the salience of the double nesting that privileges the groups Hume lists for the purpose of the construction of the steady point of view.

The upshot is that the categories of family, friends, acquaintance, etcetera, serve as the template for sympathetic response within the steady and general point of view. This entails what might be though to constitute some "formal" features of moral judgment. Every person is a member of at least one of the groups—every person is either a fellow countryman or a foreigner—so that Hume's categories are exhaustive. Further, since the groups are ordered by set inclusion, if we sympathize with family, with friends who are not family, with acquaintance who are not friends, and so forth, we will sympathize with every person exactly once. Basic moral judgements in the steady point of view result from a construction in which every person counts and there is no double-counting.

If I have Hume's picture right, the nested series of groups—family, friends, and so forth—will lessen any effects of residual bias, bias that floats free of misinformation. As we have seen, Hume's weighted, partialist universalism implies that the virtuous person will be more generous to family and friends than to strangers or indifferent persons. In arriving at these judgments through the sympathetic templates, prejudicial responses, as to the disfigurement, are screened off. In sum, it is plausible that, at the level of basic moral judgments, any distortions in sympathetic response owing to residues of bias will have a marginal impact: the doubly nested series of categories that constitutes basic moral thinking filters out potentially problematic groupings.

4. Common Morality and Roles for Reflection

Let me return to Kelly's question about the extent to which the resources of Hume's theory depend upon reflective endorsement. We have reviewed a number of elements essential to the characterization or the moral point of view:

(1) the perspective from which the moral judge sympathizes—that of the agent's narrow circle (part 2);

(2) the usual or typical circumstances associated with a character trait (part 2);[20]

(3) an informational requirement—perhaps in terms of information readily available within the narrow circle (part 3);

(4) the templates through which sympathy operates—the double nesting of family, friends, acquaintance, countrymen, foreigners (part 3).

Perhaps the members of the agent's narrow circle tend to be excessively sensitive to the feelings of others, or to be psychopaths with little or no capacity for empathy. To what extent are agents at the center of these circles subject to correspondingly stringent and permissive moral requirements, respectively? Addressing this source of variability requires a further element:

(5) the sympathetic sensitivity characteristic of the steady and general point of view.[21]

I have not mentioned this dimension until this point because I doubt that tinkering here helps address Kelly's cases.

A full specification of the steady point of view is an elaborate construction. Considerable reflection is required *to characterize* this construction. To what extent is reflection also required *to adopt* the steady point of view? Answering this question may have to proceed case by case with respect to (1–5). I will simply offer some observations about the shape of the question and Hume's answer to it.

Consider Hume's project in 3.3. He seeks to explain how the vulgar or common person arrives at moral judgments, much as 1.4.2 attempts to explain how the vulgar arrive at the belief in body (SJ 122–3). Similar issues arise in connection with Hume's explanation of how the vulgar arrive at beliefs with philosophical or metaphysical content (part 4 of my reply to Schmittt). There are also parallels in Hume's theory of justification (part 4 of my reply to Williams). The explanations, as I view matters, have normative import (part 2), but my present point is that they are explanations of the beliefs of common persons. If they are to succeed, any reflection that comes into play is of necessity fairly rudimentary. How could it be that the common person arrives in the steady point of view, absent the considerable reflection required to characterize that point of view?

A partial explanation is this: some of the specifications are matters of psychological salience in the circumstances of the uneasiness and discomfort owing to a feeling of contradiction. It is precisely in virtue of their salience that features of the steady and general point of view are stabilizing. In describing these elements of the moral point of view, a Humean theorist invokes such notions as stability, salience, double nesting, and the like. The common person, however, need not consciously reflect on these notions in arriving at the steady point of view; rather, main elements of that point of view are salient responses to uneasiness. This explains, schematically, how the common person can enter into a point of view that, in broad outline, incorporates features of the construction.

Hume is fundamentally concerned with the evaluation of character traits of agents. It seems plausible that a common person can arrive, unreflectively, at the point of view specified in (1) and (2) because focusing on an agent's narrow circle and on a character trait's usual effects are salient in the circumstances of variability and the attendant discomfort. Similarly, in connection with (3), focusing

on readily available information is salient in the face of psychological pressure owing to variability; even the common person will feel discomfort at the prospect of judgments likely to be upended by new information. In connection with (4), the groups with whom the moral judge sympathizes are salient because of the convergence of nestings of set membership and psychological strength. Hume's strategy is to rely on psychological salience to explain how the vulgar arrive in the steady and general point of view. A nice feature of this approach is that it can generate default settings for some of the elements of the moral point of view. For example, in connection with (5), perhaps the default is the "normal" degree of sympathetic sensitivity .

Of course, there is room for shortfall. Unreflective salience for the common person at best produces a vague specification of the composition of the narrow circle in (1), of the "usual circumstances" in which a character trait manifests itself in (2), of "readily available" information in (3), of the precise boundaries of the groups in (4), and of a "normal" or otherwise relevant degree of sensitivity in (5). Even where salience is engaged, it may fix an element of the steady and general point of view only in broad outline, consistent with a range of moral assessments; vulgar or common morality may be somewhat indeterminate. (This is perhaps a virtue in Hume's position, insofar as it matches or captures indeterminacy in ordinary moral thought.) To the extent that this is the case, reflection might play a role in *refining* common morality. It is entirely compatible with Hume's position to distinguish between the somewhat indeterminate steady point of view characteristic of common morality, on the one hand, and a more highly articulated conception of the steady point of view that results from reflective refinement of the common moral standpoint, on the other. But the role of reflection would be confined to the theory's penumbra. I hope to have shown that the stability conception of normativity is a generous endowment within Hume's sentimentalist ethics.

II. Reply to Schmitt

1. The Stability Account of Belief and the Aim of Truth (Schmitt, part 1)

One of my central claims is that Hume's stability account of justification is integrated with his account of belief in terms of steadiness, itself a form of stability (SJ § III.2). Schmitt agrees that, for Hume, steadiness is both a natural function of belief and a part of its nature, but holds that guidance of the will shares these characteristics; further, Schmitt thinks that the function of steadiness is at least partly derivative from that of guidance. By "guidance," he means "guidance of the will by foresight" (Schmitt 298), noting that Hume's use of the term 'foresaw' looks epistemic. A passage about foresight at 1.3.8.13, among other texts (part 3

and part 4 of my reply to Williams), supports this reading. Since Hume cannot be committed to the view that beliefs are by nature true, "[t]he best interpretation" (Schmitt 299) is that "the natural function of belief, and hence its nature, entails that most of the guiding beliefs are true representations of the future" (Schmitt 299). The next step is to suggest that successfully obtaining goods and evils is the most fundamental value and that the value of steadiness is subservient to it. This is an appealing picture.

I agree that successful guidance is a theme in 1.3.10. I do not accept the inference to the claim that it is of the nature of belief that most guiding beliefs are true. This is because I am not "[a]ssuming . . . that the natural functions of belief belong to its nature" (Schmitt 299). I set out textual evidence that steadiness is essential to belief, evidence that is independent of 1.3.10 (SJ §§ III.1–3). I do not appeal to this section's early paragraphs to establish the steadiness account of belief, but to explain why Hume focuses on stability in his epistemology (SJ § III.4): for Hume, the distinction between justified and unjustified belief tracks conditions that respectively promote and interfere with the natural function of belief. I do not take Hume to infer belief's nature from its natural function.

At the same time, the second and third paragraphs of 1.3.10 confirm my interpretation of Hume's account of the nature of belief. Schmitt quotes the final three sentences of the second paragraph. It is worth considering the beginning of the third paragraph:

> Nature has, therefore, chosen a medium, and has neither bestow'd on every idea of good and evil the power of actuating the will, nor yet has entirely excluded them from this influence. Tho' an idle fiction has no efficacy, yet . . . the ideas of those objects, which we believe either are or will be existent, produce in a lesser degree the same effect with . . . impressions. . . . The effect, then, of belief is to raise up a simple idea to an equality with our impressions, and bestow on it a like influence on the passions. (SBN 119)

Nature's "medium" is belief. When Hume writes of belief's "effect," he is identifying a defining characteristic; Hume concludes the paragraph with a definition of belief. He takes these developments to "serve as an additional argument for the present system" (SBN 120), the account of belief developed in earlier sections of part 3.

By contrast, a function of *successful* guidance, guidance by *true* beliefs, has no apparent role in the third paragraph. Hume writes of the effects of "the ideas of those objects, which we believe either are or will be existent"; "foresight" gives way to (mere) belief. Of course, one can read the theme of guidance into paragraph three. My point is that Hume draws a conclusion about what belief is—"nothing but . . . " (SBN 120)—one that contains no veritistic language.[22] Schmitt allows

that I could maintain my interpretation if I were "were able to impute one natural function, that of steady influence, to the nature of belief while denying that the other natural function, guidance, belongs to the nature of belief"; he sees "no basis, in this passage [the second paragraph] at least, for discriminating between the two natural functions in this way" (Schmitt 299). There is such a basis in the asymmetry in Hume's conclusions at paragraph three.

Suppose we grant that, for Hume, successful guidance is a natural function of belief and that the natural functions of belief belong to its nature. Suppose we also grant that this implies that most guiding beliefs are true (though there is no simple entailment here). Still, Hume does not draw this conclusion; it is not his own uptake, in the third paragraph, from the premises. Schmitt grants that I am "right that there is a good deal more talk of stability in Hume than of truth and reliability" (Schmitt 323). Section 1.3.10 contains a fascinating instance of this phenomenon. The second paragraph is welcoming to veritistic considerations, but these are a "no show" in the next paragraph when Hume gives an account of the nature of belief. A fully successful interpretation would explain why Hume gives stability more play than truth. I return to this matter in part III.

Schmitt introduces a philosophical objection to the pure stability account of belief: "It is a commonplace that belief aims at truth. It is plausible that this is so by the very nature of belief. On the face of it, a stability account is inconsistent with this claim; for it makes no mention of the aim of truth" (Schmitt 300). The point of the commonplace is that there is a second-order disposition to regulate one's belief by (what one takes to be) evidence or indicators of truth. Schmitt argues that the pure steadiness account cannot provide a rationale for the presence of this second-order regulative disposition. Since Schmitt takes belief to aim at truth "by the very nature of belief," he understands the second-order disposition as a "requirement" (Schmitt 300, 301) for belief, so that the steadiness account must provide a rationale for the presence of the disposition "in *all* cases of belief" (Schmitt 302; cf. 300).

Some philosophers do maintain that a regulative disposition is essential to belief (SJ 82n38). Is this Hume's view? I allow that, for Hume, the disposition is "closely associated with," "generally characteristic of," "typically associated with," belief (SJ 83, 84). In 1.3.10, Hume nevertheless declines the opportunity to include any property other than steadiness within the nature of belief. I am explicit that "Hume does not say whether the second-order, regulative disposition of 'aiming at truth' is . . . essential to belief" (SJ 84). Since there is no evidence that Hume sees the regulative disposition as essential to belief, the steadiness account need not provide a rationale for its presence in every case of belief.

This conclusion can be strengthened. Animals have beliefs. The point of 1.3.16 is that the animal realm provides "a strong confirmation, or rather an invincible proof of my system," of Hume's account of "that act of the mind, which we call

belief, and . . . of the principles from which it is deriv'd" (T 1.3.16.8; SBN 178). (The continuities between animals and humans in cognition, as well as belief, will prove important in part 4 of my reply to Williams.) The regulative disposition, however, would seem to require—as Hume writes in discussing "those systems, which philosophers have employ'd to account for the actions of the mind"—"a subtility and refinement of thought, as not only exceeds the capacity of mere animals, but even of children" (T 1.3.16.3; SBN 177). Hume would not think the regulative disposition present in animals, so that he is *precluded* from taking it to be essential to belief.

This result does not free the pure stability account of any explanatory burden. It must provide a rationale for a connection between belief and the regulative disposition, an explanation of why the presence of the disposition is in the service of the natural function of belief as steadiness. Rather than explain why the disposition contributes to stability in every case of belief, such a rationale would explain why it should tend to contribute to stability in the sorts of creatures that do possess the disposition. Can Schmitt's arguments against the possibility of providing a rationale for a *requirement* of a regulative disposition be adapted against providing this relaxed rationale for the disposition?

Schmitt's "first point" is that the regulative disposition is merely conditional, in that it must be triggered by suitable reflection (Schmitt 301). It will not be triggered in unreflective subjects, so that it will not contribute to actual stability for these subjects. I think it would suffice to show that the disposition, *when triggered,* contributes to stability (cf. Schmitt 302).

Schmitt's main point is that for reflective subjects in whom the regulative disposition is triggered, "the disposition has a *de*stabilizing effect" (Schmitt 301); it leads to instability. Schmitt invokes a claim of Hume's that I myself cite (Schmitt 300): "When I reflect on the natural fallibility of my judgment, I have less confidence in my opinions" (T 1.4.1.6; SBN 183). There are two strands in my reply, related to the roles of this quotation in my argument and in Hume's argument, respectively. Let us call the claim that reflection on fallibility reduces degree of confidence *the reduction thesis*. Hume maintains that we ought to control or correct belief in light of the fallibility of our cognitive faculties. Further, he insists that this correction accords with the reduction thesis, which plays a role in each of the phases of the argument of 1.4.1. First, the certainty in demonstrative knowledge degenerates into probability. Second, any probability judgment is subject to correction and hence to reduction in degree of confidence. Third, this process ought to proceed ad infinitum, leading to the "total extinction of belief and evidence," the reduction of probability "to nothing" (T 1.4.1.6; SBN 183). This conclusion renders "Of scepticism with regard to reason" notoriously suspect.

We can ask whether Schmitt's main point relies on the modestly destabilizing effects, described at the second phase, of a single correction (a reduction in confidence),

or the radically destabilizing effects of the repeated series of corrections, described at the third phase (the reduction "to nothing"). The latter is a Pyrrhonian result that proceeds from "sophisticated reflection" or "sophisticated Humean philosophical reflection" (Schmitt 302, 307). In a retrospective summary, Schmitt writes that he has urged that "*ordinary* reflection on the fallibility of one's judgment destabilizes belief" (Schmitt 307, emphasis added). This is the position he should take; the fact that the sophisticated reflection or hyper-reflection of the third phase would destabilize belief is consistent with the possibility that the sorts of reflection that more typically trigger the regulative disposition tend to enhance stability.

Although Schmitt's point that reflection on fallibility is destabilizing is not based on the Pyrrhonian result at the third phase, the reduction thesis to which Schmitt appeals is an obvious villain in determining the argument's outcome. Suppose I have a very low degree of confidence in a belief that *p*. In fact, the belief results from long-term memory, and assurance or conviction decays in proportion to the remoteness of memory (T 1.3.13.2; SBN 143–4). Suppose that I believe (justifiably) both that the belief that *p* results from long-term memory and that my long-term memory, though fallible, is highly accurate. These beliefs should engage the regulative disposition, with the result that my degree of confidence in my belief that *p* is increased; in some cases, reflection can and ought to fortify belief. The mere fact that I reflect on the (degree of the) fallibility of my faculties need not have a destabilizing effect, even in cases where the disposition is triggered. In 1.4.1, Hume misses this obvious point. Since demonstration initially involves certainty, the highest degree of confidence, any correction to demonstrative knowledge must consist in a reduction in confidence. From this perspective, Hume's insistence that the correction of any belief reduces confidence looks like a hasty generalization from the case of demonstration. (I explain Hume's error, without relying on such a crude mistake, in § VII.3 of my book.) The reduction thesis is sufficiently wrongheaded that I am reluctant to take it as evidentiary with respect to Hume's views in regard to the regulative disposition.

I call attention to the reduction thesis in order to show that Hume is committed to the presence of the regulative disposition, at least in some humans (SJ 83–4). It is difficult to see why reflection on fallibility would diminish confidence unless one were governed by the disposition; the reduction thesis presupposes the disposition. This interpretive argument does not depend upon the idiosyncratic features of the reduction thesis. Suppose Hume were to claim instead that reflection on fallibility reduces or increases confidence, as appropriate to the case. This more general—and more apt—position also presupposes that belief aims at truth; reflection on fallibility would not lead to adjustments in confidence unless one sought to regulate belief by what one takes to be evidence of truth. There is no need to let the ill-conceived reduction thesis infect an account of Hume's general posture toward the regulative disposition.

Schmitt writes of "the regulative disposition at T 1.4.1.6" (Schmitt 301, 302). This can be misleading. Reflection on fallibility does not constitute the regulative disposition; rather, it is but one of many possible kinds of reflection that can engage it. Schmitt does "not deny that there may be some second-order disposition to regulate belief with respect to truth . . . that when triggered routinely contributes positively to actual stability. . . . My point is only that the regulative disposition at T 1.4.1.6 . . . does not do this" (Schmitt 302). Correction in accordance with the reduction thesis is an operation of a more general regulative disposition. Schmitt allows that reflection upon "second-order general rules—generalizations about the success of classes of first-order generalizations—are stabilizing for Hume [SJ § IV.2]" (Schmitt 324n4). This reflection on general rules also engages the regulative disposition. Granted, such reflection is "not general across the population" (Schmitt 324n4), so that its stabilizing effects will not sustain a rationale for the *requirement* of a regulative disposition. But they might sustain an explanation of how the disposition serves the natural function of belief by tending to enhance stability.

2. Actual versus Reflective Stability (Schmitt, parts 2–3 and Williams, part 6)

Schmitt observes that "Loeb supports the stability interpretation of justified belief by appeal to the fact that Hume often makes psychological and epistemic claims in tandem; Hume passes without notice from talk of belief-formation to talk of justified belief" (Schmitt 302). Schmitt "urge[s] that we can do without the stability interpretation in one location extensively discussed by Loeb—the sixth paragraph of 1.3.9" (Schmitt 303). I capitulate; "[a] role for the stability account seems unlikely" (Schmitt 306) in this paragraph.

Let me put this concession in perspective. I bring my interpretive strategy to bear on four examples of tandem passages: (1) paragraphs three and four of 1.3.9—Hume's discussion of the two systems of realities (SJ 62–4); (2) passages in 1.3.2,4 straddling the claims that causal inference is the source of belief and that it is the source of knowledge (SJ 61, 63, 77); (3) the claim that causal inference is "just" (T 1.3.6.7; SBN 89) in close proximity to sections discussing the nature, causes, and influence of belief (SJ 61); and (4) Hume's tendency in 1.3.11–13 to identify degree of belief with degree of evidence (SJ 101–2). In addition, features of the 1.3.9 discussion of education support the stability interpretation (SJ 60, 77–9).

My half-page discussion of the sixth paragraph of 1.3.9 (SJ 63) is one of four points (SJ 62–4) intended to resist an alternative reading of passage (1), to protect a flank. (At page 73, I do treat the discussion of resemblance as a direct argument for the stability interpretation). Schmitt is prepared to "concede that [the 'tandem' references to belief and justified belief] cumulatively provide prima facie evidence for the stability interpretation" (Schmitt 303); he raises no direct objection to my

extracting the stability account of justification from the tandem passages (1–4). Schmitt's discussion of the sixth paragraph of 1.3.9 leaves standing "my most important single line of argument" (SJ 31) for interpreting Hume as offering a stability account of justification.

The tandem passages support attributing to Hume a theory of justification in terms of stability, but they do not favor any particular version of a stability account (SJ 90–1). As Schmitt's discussion makes plain (see the next part), I am less than resourceful in considering options available to a stability theorist. I do distinguish two versions of a stability account that might be attributed to Hume (SJ 87–9). Schmitt's characterizations (Schmitt 306) are fair enough, provided we have a fix on my terminology. On the *reflective* stability account, a belief is justified for any subject (however reflective) if it results from a mechanism that would tend to produce stable beliefs were it used by a reflective person. On the *actual* stability account, a belief is justified for a subject if it results from a mechanism that tends to produce stable beliefs given the actual degree to which that subject is reflective. The terminological point is that I characterize a "(fully) reflective person" as one "who searches for contradictions among beliefs, examines the reliability of belief-forming mechanisms in producing true beliefs, considers the results of methodical application of the cognitive faculties, and so forth" (SJ 88).[23] My intent is that a (fully) reflective person is given to intense reflection, the sorts of sophisticated reflection that generates Pyrrhonian doubt.

I refer to persons who are not (fully) reflective as "unreflective." This terminology is misleading: short of intense reflection, there are many degrees of reflectiveness. Greater reflection tends to produce greater stability, so that there are gradations in epistemic success among those who are not hyper-reflective. For example, the beliefs of persons who do not reflect on readily available considerations and evidence will be vulnerable to destabilization.[24] "Unreflective" persons, even though not (fully) reflective, can be highly reflective. It would have been more perspicuous to refer to (fully) reflective persons as "intensely reflective" or "hyper-reflective."[25] To help avoid misunderstanding, I will sometimes employ scare-quotes—"unreflective"—in referring to persons who are less than fully reflective.

In distinguishing the actual stability and reflective stability accounts, I seek to emphasize that their implications diverge in the case of subjects who are not fully or hyper-reflective (cf. SJ 89). On an actual stability account, the destabilizing results of intense reflection are of no relevance to the epistemic status of "unreflective" subjects. By contrast, a reflective stability account implies a lower epistemic standing for the beliefs of "unreflective" subjects than does an actual stability account. I provide three reasons for attributing the actual stability account to Hume: a philosophical reason, rooted in considerations of the value of stability; a large-scale interpretive reason, rooted in Hume's project of systematizing his

pretheoretical epistemic commitments; and a local textual reason, rooted in the honest gentlemen passage (SJ § III.6). I take these up in reverse order.

I maintain that Hume's discussion of the honest gentlemen (T 1.4.7.14; SBN 272–3) has an "approving ring," or at least that "Hume is unwilling to criticize" them; since the reflective stability account would detract from the justification of the landholders, I take this to support an actual stability account (SJ 92–3, 96–7). Janet Broughton has suggested, in conversation, that Hume is simply condescending to the country gentlemen. Williams makes the further point that I overlook the use to which Hume puts the discussion in the continuation of the passage beyond the portion I cite (Williams 32–5). Although I do not think Hume so sneering toward the landholders as Williams suggests (part 4 of my reply to Williams), he is right that my treatment of this material is flat-footed. Fortunately, I appeal to the country gentlemen passage in order to confirm an interpretation for which I advance other arguments.

I take the constructive stage of the development of Book 1 to reflect Hume's pretheoretical intuitions about justified belief. Beyond that, I claim that Hume's goal in the constructive stage is to put forward a theory of justification that systematizes and explains, and thereby sustains, his pretheoretical epistemic assessments (SJ 12–14, 32, 59).[26] Consider, for example, Hume's highly positive evaluations of beliefs based on causal inference (Schmitt 302). Intense reflection is destabilizing; if Hume subscribes to the reflective stability account, he must lower, if not discard, all his positive assessments. If Hume subscribes to the actual stability account, he need only lower or discard the positive assessments that apply to the beliefs of an intensely reflective person. Since causal inference is not unstable and hence not unjustifying for "unreflective" persons, Hume can fully retain his positive assessments of their beliefs based on causal inference. By contrast, the reflective stability account requires systematic revision of those assessments. This provides a decided reason to prefer an interpretation in terms of the actual stability account. This *ratification argument* depends upon the differential capacities of the actual and reflective stability accounts to systematize and thus ratify Hume's pretheoretical distinctions. This is the heart of my large-scale interpretive argument (SJ 91–2).[27]

I note a connection between the ratification argument and my intended use of the honest gentlemen passage. I rely on that passage to conclude that, even in the face of the destabilizing effects of intense reflection in 1.4.7, Hume does not relinquish or modify any positive evaluation of the beliefs of the landholders.[28] This is the result the ratification argument predicts; it applies to the landholders as a special case of persons who are not intensely reflective. Williams of course rejects the ratification argument. He grants that Hume must rely on pretheoretical intuitions about justified and unjustified beliefs or inferences, but maintains that "a systematic account, and thus a kind of vindication, of his epistemic intuitions . . . is not Hume's goal at all"; rather, "the point of Hume's investigations" is reform

of the sciences (Williams 284). I defer consideration of this alternative interpretation until the final part of my reply to Williams.

Schmitt's formulation of the ratification argument differs from my own; he focuses on the thought that Hume's claim in part 3 that causal inference is justified is "logically inconsistent" (Schmitt 308) with his claim in 1.4.7 that causal inference yields unstable beliefs when used by a reflective person. I emphasize that on the reflective stability account the epistemic evaluations in part 3 are indeed provisional (cf. Schmitt 309) and subsequently modified in light of new findings (SJ 88, 90–2). We can grant that the positive evaluations in part 3 are intended to apply to all subjects, reflective and "unreflective" (Schmitt 308–9). In virtue of the provisionalist character of the reflective stability account, this is harmless: the differing assessments in parts 3 and 4 are indeed "reconciled by some nonlogical device" (Schmitt 309); they reflect the psychology and philosophy available at different stages of inquiry. Once the instabilities are uncovered in 1.4.7, Hume systematically discards or lowers earlier, pretheoretical assessments in regard to all subjects, reflective and "unreflective." The advantage of the actual stability account is that, in light of 1.4.7, Hume need alter only those of his positive evaluations that apply to intensely reflective subjects. The essential idea behind the ratification argument is not that the reflective stability interpretation cannot remove a logical inconsistency, but that Hume can better preserve his pretheoretical commitments in part 3 if he opts for the actual stability account. Schmitt transmutes an argument about *preserving* as many pretheoretical intuitions as possible into one about *reconciling* differing evaluations in parts 3 and 4.[29]

Whereas Schmitt takes the ratification argument to be my "main" reason for preferring the actual stability account (Schmitt 308), I offer it as "a useful starting point" (SJ 91) for adjudicating between the two versions of a stability theory. I discuss the argument based on the value of stability at pages 97–8 (cf. SJ 27–8, 90n48). Its basic idea is that insofar as intense reflection is destabilizing, the reflective stability account finds no rationale in an epistemology that prizes stability. Schmitt deems this objection "basically right" (Schmitt 307). One of his qualifications, that intense reflection may be instrumental to stability, converges with a main thrust of Williams's contribution. Bracketing this issue, which I take up in part 3 of my reply to Williams, the argument based on the value of stability remains intact. Both the ratification argument and the value of stability provide important considerations against the reflective stability account.

3. Peak Stability versus Average or Temporal Stability (Schmitt, parts 2–3)

Unfortunately, my book obscures the fact that both the actual stability and reflective stability accounts are best viewed as representing classes of theories

with significantly different properties. These accounts specify the highest level of reflection relevant to the epistemic evaluation of a subject's beliefs. In the reflective stability account, this level is that of intense reflection; in the actual stability account, it is "the most reflective level of investigation the actual subject has achieved" (SJ 100n61). During many periods, a subject will operate at a level of reflection below the highest level he ever achieves. In my book, I nevertheless suppose that on either the actual or reflective stability account, a subject who achieves intense reflection during some periods has no justified beliefs during any period. (See note 27 above.) I surreptitiously construe the two accounts in such a way that the sole determinant of the status of beliefs, even beliefs held during periods of lesser reflection, is how they would fare at the highest level of reflection the account brings into play. These *peak reflection accounts* submit every belief to the simple test of stability under this highest level of reflection.

In raising "a doubt about whether either stability account, the actual or the reflective, has a rationale in the value of the stability" (Schmitt 309), Schmitt has peak reflection accounts in view. He calls attention to alternatives: "whether an *unreflective belief of a reflective person* is justified should turn on whether the mechanism that produces it yields stable beliefs during the *unreflective moments* of that person's life" (emphases added). I construe "turn on" to mean "turn on, at least in part." Let us call this principle about the evaluation of beliefs held during unreflective periods *requirement (U)*. At least two variants of each of the actual stability and reflective stability accounts satisfy (U). On an *average stability* account: a justifying mechanism must tend to yield high average stability over its total output in the subject (cf. Schmitt 309–10). On a *temporal stability* account, a justifying mechanism must tend to yield stability over the temporal portion of the subject's belief system to which a belief under evaluation belongs (cf. Schmitt 324n7). Neither the average nor the temporal stability variants of either the actual stability or reflective stability accounts implies that a subject who achieves intense reflection never has justified beliefs. (See part 1 of my reply to Williams.) I thus overlook important alternatives to peak reflection accounts.[30]

Theories that satisfy requirement (U) have considerable appeal. The ratification argument favors an actual stability account because it better preserves Hume's pretheoretical commitments in regard to the epistemic status of beliefs of persons who are not intensely reflective than does the corresponding version of a reflective stability account. (For example, comparing peak reflection versions of the two accounts, the actual stability account does better.) Likewise, the ratification argument favors average or temporal stability variants of an actual stability account; they better preserve Hume's pretheoretical commitments in application to persons who achieve intense reflection than does a peak reflection variant. The argument from the value of stability favors actual stability accounts because there is no motivation for assessing beliefs in terms of a destabilizing level of reflection the subject never

achieves. Likewise, the value of stability favors the average or temporal stability variants; in the case of a subject who does achieve intense reflection, there is no motivation for assessing all beliefs in terms of a destabilizing level of reflection the subject only occasionally achieves. If stability is what we value, destabilizing intense reflection should not receive all the weight in epistemic assessment, as it does in a peak reflection version of an actual stability account applied to an intensely reflective subject. Schmitt is correct: "a well-motivated stability account" ought to satisfy requirement (U) (Schmitt 324n8; cf. 309–10).

Although Schmitt considers requirement (U) well-motivated, he raises a textual objection to the average stability variant: on the average stability account causal inference is justified for intensely reflective subjects since they are relatively unreflective most of the time, but 1.4.7 gives a negative evaluation of causal inference for these subjects (Schmitt 308–9, 310). Schmitt allows that the scope of the epistemic evaluations in the Conclusion is a "very broad and difficult interpretive issue" (Schmitt 308). There is room to suggest that when Hume is "ready to reject all belief and reasoning," so that he "can look upon no opinion even as more probable or likely than another" (T 1.4.7.8; SBN 268–9), this readiness—within an "*intense* view" (T 1.4.7.8; SBN 268)—is exaggerated and need not represent a negative evaluation of intensely reflective subjects tout court. This approach weakens Schmitt's argument.

Schmitt generalizes his objection to the average stability version to any interpretation that satisfies requirement (U): "any stability account motivated by the value of stability . . . will be inconsistent with Hume's negative evaluation of causal inference in reflection" (Schmitt 310). This generalization loses sight of the temporal stability account (which Schmitt relegates to a footnote, 324n7). From its perspective, the negative evaluation of causal inference applies to reflective subjects only during their reflective periods; the temporal stability account, unlike the average stability account, does not even appear at odds with the textual phenomena in 1.4.7.

Let me summarize this and the preceding part. I assume the success of the main argument extracting a stability account from the tandem passages. The ratification argument and the argument from the value of stability provide reasons to prefer an actual to a reflective stability account, but also to prefer an average or temporal stability variant to a peak reflection account. Both these variants can survive Schmitt's objections. At the same time, I leave open the comparative merit of the average and temporal stability interpretations, all things considered. The case for average stability might in part depend upon the frequency, duration, and stability of the relatively unreflective periods. One could debate whether the ratification argument or the value of stability provide resources for favoring one of these variants over another (cf. SJ 100–1). What is clear is that in my book I ought to have advanced an interpretation satisfying requirement (U), rather than the peak reflection account.

4. Metaphysical Beliefs and the Identity Propensity (Schmitt, parts 4–5)

Schmitt offers a number of criticisms of the stability interpretation insofar as it relates to the belief in material substance and the philosopher's belief in body. He is "uncertain that in Hume's sampling of the philosophical beliefs that result from the identity propensity . . . there are enough unstable beliefs to make the propensity unjustifying according to the stability account" (Schmitt 315–16). He also suggests that the actual stability account cannot accommodate the role of reflection in producing these beliefs. Finally, Schmitt claims that the stability interpretation cannot account for Hume's method of evaluating the identity propensity itself.

Schmitt questions whether Hume takes the belief in material substance or substratum to be unjustified because it is unstable (Schmitt 311). There is no question that Hume invokes the identity propensity at paragraphs three through five of 1.4.3 to explain the belief in substance (Schmitt 311). Schmitt takes the view that this material constitutes "a *purely* psychological description" (Schmitt 311, emphasis added) of the origin of the belief, at least if we put aside an occurrence of a veritistic term, "deceives" (T 1.4.3.3; SBN 220).

Schmitt interprets Hume's argument as proceeding in two main parts (Schmitt 311–12), which he finds in the seventh paragraph of 1.4.3 (SBN 222):

First part:

(1) "Every quality . . . may be conceived to exist apart . . . from every other quality."

(2) Therefore, the belief in accidents is unreasonable.

Second part:

(3) The belief in accidents is entailed by the belief in material substance.

(4) Therefore—from (2) and (3)—the belief in material substance is unreasonable.

Schmitt considers various reconstructions of the transitions in the two parts (Schmitt 312–14) and concludes that the texts are incompatible with a role for stability. According to Schmitt, Hume "devotes most of his attention to arguing that the ancient belief in accidents is unreasonable" (Schmitt 311–12).

I read the seventh paragraph (SBN 222) as follows:

(a) The belief in accidents is entailed by the belief in material substance.

Step (a) is identical to Schmitt's (3). (Strictly, Hume says that the belief in substance has "the notion of" accidents for an unavoidable consequence.) The next claim is as follows:

(b) We also believe that accidents "cannot subsist apart, but require a subject of inhesion to sustain and support them"; that is, we believe in "a dependance of every quality on the unknown substance."

Hume distinguishes the belief in the existence of accidents and the belief in their *dependence* upon substance. Immediately after writing that belief in accidents is a consequence of the belief in material substance, he adds: "*nor can we forbear* looking upon colours, sounds, tastes, figures, and other properties of bodies, as existences, which cannot subsist apart, but require a subject of inhesion" (emphasis added). Although Hume devotes most of the paragraph to the belief in the dependence of accidents on substance, this belief plays no role in Schmitt's two-part argument.

The second and third sentences provide a (debunking) explanation of the belief at (b):

(c) The belief in the dependence of accidents upon substance arises from custom.

The basic idea in support of (c) is that whenever we observe sensible qualities, we also, "for the reasons above-mention'd, . . . fancy a substance to exist"; habit then leads us to infer the dependence of the qualities on the substance. I do not see what "the reasons above-mention'd" could be, if not the psychological explanation of the belief in substance at paragraphs three through five. The first sentence of paragraph seven notes that belief in accidents is an unavoidable consequence "of this method of thinking with regard to substances," a way of thinking explained in previous paragraphs.

Following the explanation of the belief in the dependence of accidents, Hume writes in the fourth sentence that "[t]his conceit" is "unreasonable." In the fifth and final sentence of the paragraph, he offers a conceivability argument for this result:

(d) "Every quality being a distinct thing from another, may be conceiv'd to exist apart, not only from every other quality, but from . . . substance."

Schmitt's partial quotation at (1) omits Hume's claim that qualities are conceivable apart "from substance." This claim is essential if Hume is to bring his argument to bear on the conceit of dependence. Hume thus concludes:

(e) The belief in the dependence of accidents upon substance is unreasonable.

There is no conclusion here along the lines of Schmitt's (4). In the seventh paragraph, Hume presents no argument against the reasonableness of the belief in the *existence* of accidents. Even if the belief in substance entails the belief in the existence of accidents, it does not—as Hume presents matters—entail the belief in the dependence of accidents upon substance. This does not settle the question of the best interpretation of the earlier paragraphs, but paragraph seven does not undermine the stability interpretation of Hume's account of the epistemic status of the belief in substance.

I turn to the belief in the double existence of impressions and objects. Schmitt doubts that the philosophical belief in body is unstable. Even if Hume

does not remark that this belief is capricious (Schmitt 314), in the penultimate, summary paragraph of 1.4.2 he does attribute the belief to the "trivial qualities of the fancy" (SBN 217). This language has affinities with the final paragraph of 1.4.3: the ancients "have shown they were guided by every trivial propensity of the imagination" (SBN 224; cf. Schmitt, 325n11). Schmitt's principal strategy is to undermine textual evidence that "may suggest that the philosophical belief in body is unstable." Schmitt mentions, but does not discuss, Hume's claim that the philosophical belief in body is "palliative" (cf. Schmitt 314)—"only a palliative remedy" (T 1.4.2.46; SBN 211).

The philosophical belief is merely palliative because it is a *conflicted* resolution to an underlying contradiction between the vulgar belief and reflection (SJ 152–4). To see this, consider that the identity propensity induces the vulgar to ascribe a continued existence to the very impressions that they perceive. Reflection on double vision and perceptual relativity shows that impressions do not have a continued existence (T 1.4.2.25; SBN 210–1). In ascribing continued existence to a second set of objects, to objects other than observed impressions, the philosophical belief goes some way toward satisfying the inclination to ascribe identity to the interrupted impressions. Yet, belief in the continued existence of a second set of objects cannot fully satisfy the propensity. This belief *mislocates* the continued existence that the vulgar are inclined to ascribe to observed impressions, placing it instead in wholly unobserved objects. The philosophical belief merely "seems" to comprehend both the identity propensity and reflection; it is a new fiction by which we "elude" the contradiction between these principles, a "pretext" for maintaining both the vulgar belief and the results of reflection (T 1.4.2.52; SBN 215, 216). Schmitt himself recognizes the conflicted character of the belief, which "*attempts* to honor . . . reflection" (Schmitt 314, emphasis added). Even granting that the philosophical belief is unstable, Schmitt could contend that Hume does not trace the belief's unjustifiedness to its instability. (He mentions one text that he takes to trace the unjustifiedness to veritistic considerations.[31]) The debate would have to be joined at this level.

In sum, for all Schmitt has shown, there are sufficient unstable beliefs to render the identity propensity unjustifying. Hume does maintain that the philosophical belief in body is unstable. And Schmitt's argument against Hume's tracing the unjustifiedness of the belief in substance to instability does not succeed. It is worth noting that the belief in immaterial substrata or souls inherits the instability in the belief in material substance (SJ 148–9, 151–2).

Schmitt also raises a difficulty for my explanation of the unjustifiedness of the vulgar belief in body, insofar as it is based on the constancy of perceptions. The vulgar belief is stable for the "unreflective" person, who does not reflect on double vision and perceptual relativity (cf. Schmitt 314; SJ 150–2, 161). On my interpretation, the vulgar belief is unjustified because it nevertheless results from a propensity that tends to produce unstable outputs; the propensity also produces

the beliefs in material and immaterial substance, which are unstable even for the "unreflective" person (Schmitt 316; SJ 161–2).[32] Schmitt finds a serious problem, on the ground that the beliefs in substance require philosophical reflection and hence do not belong in the hopper if Hume is evaluating the stability of the identity propensity's outputs within vulgar consciousness. Schmitt introduces "philosophically important beliefs" by enumeration—the vulgar belief in body, and the philosophical beliefs in body and in material and immaterial substances (Schmitt 21). Schmitt's label, "philosophical beliefs" (Schmitt 311, 313, 315–16), is suggestive of beliefs that arise from reflection. I prefer "metaphysical beliefs" (SJ 10, 35, 141). Do these beliefs depend upon reflection?

The identity propensity plays an essential role in a *pattern* of psychological explanation (SJ 35, 139–44, 176). (1) We ascribe identity to related objects that are interrupted or variable; this is due to the propensity to mistake resemblance for identity. (2) The interruptions or changes in the related objects induce us to regard them as different or diverse; this is because we regard identity, in a strict sense, as requiring an object that is uninterrupted and invariable. We are thus involved in a contradiction. (3) In an attempt to resolve or remove the contradiction, we suppose that there exists an object that is strictly identical, uninterrupted and invariable. The explanations of the vulgar belief in body and of the beliefs in material and immaterial substance are instances of this pattern.

Hume's description of the three-stage reaction does invoke refined concepts, for example, that of strict identity. The psychological reaction, however, does not require the subject consciously to reflect on this notion; Hume's claim is only that our psychology engages it, whether we realize it or not. In explaining the vulgar belief in body, Hume is not claiming that the vulgar consciously appeal to claims along the lines of (1–2) as steps in an argument for the belief at (3) that perceptions have a continued existence (SJ 162). Hume is emphatic that the vulgar do not rest the belief in body upon argument (T 1.4.2.14; SBN 193). The three-stage reaction unfolds at a nonconscious and nonreflective level (cf. SJ 149n6).

The operation of the identity propensity to produce the vulgar belief in body cannot involve philosophical reflection. Hume is discussing mechanisms that operate in "the unthinking and unphilosophical part of mankind, (that is, all of us, at one time or other)" (T 1.4.2.36; SBN 205). If the propensity required reflection, Hume would have failed in his objective of providing an explanation of the *vulgar* belief—an ambition that Hume applies specifically to the belief in body insofar as it arises from constancy (T 1.4.2.31; SBN 201–2). In describing the reaction that generates the vulgar belief, Hume writes of psychological effects on "us," that "we" undergo, that effect "the mind" or the "imagination" (T 1.4.2.24–43; SBN 199–210).

Interestingly, the 1.4.3 explanation of how the identity propensity leads to the belief invokes the same language—"us," "we," "the mind," the "imagination"

(T 1.4.3.2–5; SBN 219–21). In explaining the belief in substance, paragraphs two through five make no reference to psychological operations distinctive to the ancient *philosopher*. By contrast, Hume writes in the opening sentence of the sixth paragraph that "The peripatetic philosophy asserts the *original* matter to be perfectly homogeneous in all bodies" (T 1.4.3.6; SBN 221). Hume writes in the next sentence: "At the same time it [the peripatetic philosophy] assigns to each of these species of objects a distinct *substantial form*" (SBN 221). The notion of occult qualities, by which "these philosophers carry their fictions still farther" (SBN 222), makes its debut in the eighth paragraph. The doctrines of a perfectly homogenous matter, of substantial forms, and of occult qualities are elaborations of a vulgar supposition of substance, somewhat as the philosophical belief in double existence is rooted in the vulgar belief in body. The ancient philosophy has "a very intimate connexion with the principles of human nature" (T 1.4.3.1; SBN 219) because it takes for its starting point a belief in material substance that arises, via the identity propensity, without philosophical reflection. It is only the peripatetic accretions to the belief in substance that depend upon reflection. The vulgar belief in body, the belief in material substance, and the belief in immaterial substance arise within unreflective, vulgar consciousness.

I turn to Schmitt's questions about Hume's evaluation of the identity propensity. In the opening paragraph of 1.4.4, Hume raises an objection, that he is "unjust in blaming the antient philosophers for making use of" of the imagination (SBN 225). According to Schmitt, Hume commits himself to a "key constraint" or principle in the course of this paragraph: causal inference plays a "central" role in evaluating the identity propensity (Schmitt 318) and, presumably, in evaluating other belief-forming mechanisms. Schmitt maintains that a stability interpretation of Hume's theory of justification cannot account for the *objection* Hume is raising; this is because causal inference does not play a central role in evaluating the identity propensity (Schmitt 317–21).

It is unclear how Schmitt extracts the key principle. Hume does not preclude the existence of "permanent, irresistable, and universal" (PIU) belief-forming mechanisms other than causal inference. Were causal inference the sole such principle, Hume could *identify* causal inference with the PIU principles, rather than cite it as an *example* of these principles, as he does (T 1.4.4.1; SBN 225). The PIU principles include intuition, demonstration, perception, and memory, as well as causal inference (SJ 46–47, 67, 159, 248). This set of resources is central to the evaluation of the identity propensity. The imagination serves as "the ultimate judge" (T 1.4.4.1; SBN 219), but different PIU principles will be especially relevant in different cases.

Schmitt further maintains that a stability interpretation cannot account for Hume's *reply* to the objection in the opening paragraph of 1.4.1 (Schmitt 321–2). I focus on his main reason for this, his "second" point" (Schmitt 323). Schmitt

maintains that "given the context, the contrast to which Hume refers in speaking of subversion is . . . between the use of causal inference to evaluate whether the identity propensity is justifying, and the use of the identity propensity to evaluate whether that same propensity, the identity propensity, is justifying" (Schmitt 323). He then objects that "the propensity does not even judge its own stability" (Schmitt 323). That looks right, suggesting that Schmitt's sense of the context is off the mark. The relevant question is whether the set of PIU principles—intuition, demonstration, perception, memory, causal inference—yields a different assessment of the identity propensity than does the set of non-PIU principles. The latter presumably includes, in addition to the identity propensity, the propensity to add a new relation to related objects (cf. Schmitt 325n11), apprehension (as at 1.4.4.1; SBN 226), and any other belief-forming mechanism not included in the PIU principles—superstition, crystal-ball gazing, wishful-thinking.

Unsurprisingly, Schmitt maintains that a reliability account better fits the opening paragraph of 1.4.4 than does a stability interpretation (Schmitt 321): Hume evaluates the identity propensity as unjustified on the ground that it leads to false beliefs. The identity propensity plays a fundamental role in generating both the vulgar and the philosophical belief in body. Schmitt contends that Hume appeals to causal inference itself—in the double vision experiment and in the 1.4.4 argument about primary qualities and matter (1.4.4.5–14; SBN 227–31)—to show the falsehood of these respective beliefs. (I have misgivings about the role of causal inference. See note 42.) Note, however, that Schmitt cites no argument grounded in causal inference to show that the beliefs in material and immaterial substance are false.[33] These examples cannot be dismissed as peripheral to Hume's concerns. The objection at the beginning of 1.4.4 is precisely that Hume is unjust in blaming the ancient philosophers for relying on the imagination, and thus unjust in criticizing "their fictions of substance and accident" (T 1.4.4.2; SBN 226). Neither the vulgar nor the philosophical belief in body is directly on the table. The evaluation of the beliefs that are on the table are not shown false by casual inference, so that the inference to unreliability is not available.

Schmitt's reading of 1.4.4.1 is one of the many plugs he puts in for the reliabilist interpretation of Hume that he has productively advanced elsewhere.[34] Another concerns 1.4.2. The vulgar belief arises from two mechanisms: in cases of constancy, from the identity propensity; in cases of coherence, from "custom-and-galley" (SJ 35–6, 139–40, 177–80, 186–93). Schmitt maintains that a reliability account explains not only why the vulgar belief is unjustified, but also why vulgar beliefs arising from the identity propensity are justified to a lesser degree than those arising from custom-and-galley (Schmitt 316–17). This is a nifty idea. Hume, however, has in view a deeper asymmetry between these mechanisms (SJ 200–1). In the course of the discussion of constancy, Hume refers to the belief in the existence of perceptions that are not perceived as a "fiction" (T 1.4.2.36,43; SBN 205, 209). This

term is functioning pejoratively: the philosophical belief in a second set of objects is also a "fiction" (T 1.4.2.52; SBN 215), as are the beliefs in substances, accidents, substantial forms, and occult qualities (T 1.4.3.1,8, 1.4.4.2; SBN 219,224,226) and in immaterial substance (T 1.4.6.6,7; SBN 254, 255). By contrast, the term 'fiction' does not occur in Hume's discussion of custom-and-galley. To my ear, this is a signal that Hume sees the asymmetry in the epistemic status of custom-and-galley and the identity propensity as other than a difference in degree of reliability.

III. Hume's Mix of Stability and Truth

Skirmishes about particular passages and examples are unlikely to settle issues between the veritistic and stability interpretations of Hume's theory of justification. Hume is not systematic in setting out an epistemological position, in part because his interest in epistemology is secondary to his associationist ambitions (SJ 29–30, 58–9). It is too much to expect that one interpretation will be cleanly established. We can hope to make an all things considered judgment, informed by a rich sense of the textual constraints and the trade-offs among competing interpretations. I cannot undertake this project here. (Nor is this the project of the book. See SJ viii.) Instead, I propose to highlight some plausibility considerations that favor a stability account over a reliability account and related interpretations.

Schmitt allows that "there is a good deal more talk of stability in Hume than of truth and reliability" (Schmitt 323). In part 1 of my reply to Schmitt, I discuss one example, the third paragraph of 1.3.10. Similarly, Schmitt devotes a footnote to explaining why we find Hume "adverting to the permanence, irresistibility, universality, and subversive power of causal inference" (Schmitt 326n19), rather than to veritistic notions, in the first paragraph of 1.4.4.[35] There are other puzzling texts, for example: "We may draw inferences from the coherence of our perceptions, whether they be true or false" (T 1.3.5.2; SBN 84). (Williams puts considerable weight on this passage. See below in this part.) Here, there appears to be a curious indifference to truth. Can we explain the scanty textual basis for veritistic readings?

One factor receives attention in the book (SJ 22–3). Epistemic judgments are a species of normative judgment. In Hume's view, the normativity of moral judgment derives from motivation, actuating pain or pleasure. Moral judgments are appropriately corrected reports or expressions of either a particular kind of pleasure or satisfaction or of pain or uneasiness. Moral judgments or distinctions are thus "founded" (T 3.1.1.23,27, 3.2.2.11, 3.2.8.8; SBN 466, 470, 491, 546) in pleasure and pain. We might expect that epistemic evaluations are also founded in particular kinds of satisfaction and uneasiness. One of my criticisms of the Kemp Smith tradition of interpretation is that its stress on irresistibility or unavoidability cannot do justice to the sense in which we are subject to positive epistemic obligations. (In his part 4, Williams has a lovely development of this point.) On the

stability interpretation, epistemic obligation is naturalized as deriving from the motivational force of the felt uneasiness in unstable doxastic conditions (SJ 22). I want to be cautious in handling the analogy, as Hume does not develop a theory of justification in a way that parallels his theory of moral obligation.[36] Uneasiness and its motivational efficacy nevertheless play a crucial role in both contexts. Even if a reliability interpretation can find a way to accommodate this role, a stability interpretation does so much more directly.

A second factor, though speculative, helps to explain the first. On a reliability theory, justification depends upon whether, in fact, a belief-forming mechanism tends to produce a sufficient proportion of true beliefs. There is no requirement that, for a person to possess justified belief, he must have reasons to believe that his beliefs result from such a mechanism. A philosopher advancing reliabilism might nevertheless want to have the strongest grounds for confidence that important belief-forming mechanisms—such as causal inference, or superstition—are, or are not, reliable. Absent such confidence, the epistemologist would be uncertain of his theory's adequacy in generating implications that match our judgments about justification. What considerations could give Hume confidence in the central case of the reliability of causal inference?

Consider Hume's historical position. Many of his predecessors had explicitly or implicitly appealed to Divine benevolence or veracity. Hume was far removed in philosophical space from such views. Also, he was a good century too early to appeal to evolutionary explanations of reliability. Hume's explaining the reliability of the mechanism that underpins causal inference, custom or habit, would seem to require some other theory—some non-theological and non-evolutionary theory—of the reliability of custom. Hume, however, was exceedingly cautious in his willingness to venture theoretical explanations; he absorbed a Newtonian reserve with respect to "hypotheses" (T Intro.8, 1.1.4.6, 1.3.5.2; SBN xvii, 12–13, 84, Abs. 2; SBN 646, EHU 4.1.12; SBN 30).

The final paragraph of 1.3.16 bears witness to Hume's caution. Hume is dubious that any "ultimate reason" can be offered for the operation of custom; a fortiori, he does not have an explanation of why the "train of ideas" (T 1.3.16.9; SBN 179), the train of beliefs, produced by custom is reliable in its representation of the world. Indeed, "we are ignorant of those powers and forces, on which this regular course and succession of objects totally depends" (EHU 5.2.22; SBN 55); we cannot explain either the train of ideas or the course of nature, much less the correspondence between them. In the penultimate paragraph of *Enquiry* 5, the "pre-established harmony between the course of nature and the succession of our ideas" feeds the "wonder and admiration" of those "who delight in the discovery and contemplation of *final causes*" (EHU 5.2.21; SBN 54–5). For his part, Hume has no theory of the world—either of the operation of custom, on the one hand, or of the course of nature, on the other—that grounds his confidence in the reliability

of causal inference. This, I speculate, would be an obstacle to his putting forward a reliability theory of justification.

At the same time, there is much protoevolutionary thinking in Hume. This vein in his thought makes it tempting to develop Hume in the direction of a reliability, adaptivist, or proper function account of justification (cf. SJ 80n34 and Schmitt 324n3). I think it fine to say that Hume was groping in these directions. As a matter of broad intellectual temperament, any of these theories would have been congenial to Hume, absent his theoretical caution. These possibilities, together with the stability account of Hume's position, constitute a cluster of interrelated—because externalist—interpretations.[37] *We* can think of the uneasiness in instability as an evolutionary mechanism for signaling and helping to rectify mechanisms that are not reliable, not adaptive, or not in accord with proper functioning (cf. SJ 125–6n36). Hume could not articulate the evolutionary considerations supportive of a reliability theory and integral to adaptivist and (some) proper function theories. Nor was he in a position to elaborate an evolutionary explanation of uneasiness. Hume was in a position to observe the sorts of discomfort and uneasiness that infect belief, to see that they have an important role in motivating belief revision, to offer generalizations about these phenomena, and to fashion a theory of justification on this rigging. The stability theory, unlike its cousins, was thus a live option for this keen observer of mental life.

If Schmitt's interpretation assigns too prominent a role to truth, in Williams's interpretation truth recedes too far into the background. Williams observes that his interpretation of Hume's account of justification as consensus in the long term is reminiscent of Peirce: truth is what will be believed in the end of inquiry.[38] The Peircian position, however, would be incompatible with Hume's anti-Rationalism: "If truth were defined in terms of long-term consensus, stability-conducive principles of belief-formation would turn out to be truth-conducive after all. Hume never shows the slightest inclination to make such a claim" (Williams 293). No inclination to *define* truth in terms of consensus, coherence, or stability. But in calling attention to a "correspondence," "a kind of pre-established harmony between the course of nature and the succession of our ideas," Hume does make the claim that inductive inference—the central stability-conducive principle—is truth-conducive (EHU 5.2.21; SBN 55, cf. 5.2.22; 55).

These observations are germane to Williams's suggestion in regard to "the true significance of Hume's apparent love of paradox," a "tendency to insist on the ultimate contingency of things—including the workings of our own minds" (Williams 293). It is unclear that custom, for example, is "utterly contingent" (Williams 293), if that means without further explanation. Hume insists that we are ignorant of such explanations, but does not rule them out. One could try to read the final paragraph of *Treatise* 1.3.16 and the penultimate paragraph of *Enquiry* 5 as ironic; but it would be uncharacteristically dogmatic for Hume to take the view that there

is no explanation to be had. An alternative reading draws on one of Williams's themes, Hume's insistence on "sticking close to experience" (Williams 285; cf. 287). At this level of generality, the Baconian injunction to confine ourselves to "immediate causes" (Williams 289) is of a piece with Newtonian reserve about "hypotheses" noted above. In light of the proto-evolutionary thinking in Hume, it remains implausible that he would rule evolutionary theory out of court.

Is Hume's allowing that stability-conducive principles are truth-conducive consistent with his anti-Rationalism, on which Williams properly insists? It is consistent if Reason is restricted to a faculty "which enables us to grasp necessary truths and logical connections" (Williams 268). What if anti-Rationalism is the thesis that our beliefs "are not supportable by (what we can show to be) truth-conducive reasons" (Williams 266), that our "epistemic procedures . . . can be seen to be truth-conducive" (Williams 276)? The answer depends upon what it is to "show" or to "see" that causal inference is truth-conducive. An evolutionary explanation of the general success of causal inference would be a showing that relies on our best inductive practices. This is objectionable only from a perspective in which Hume starts out as a wide, theoretical skeptic about induction; Williams aligns himself with naturalist interpretations that reject this traditional reading (cf. Williams, parts 3–4).

What if Rationalism is the thesis that "our beliefs never amount to knowledge" (Williams 266), a thesis in the grip of a conception of a "Reason [that] aims at truth" (Williams 276)? I see little evidence that Hume is an anti-Rationalist in this sense. Williams's comment that "Hume may not always clearly distinguish [truth] from demonstrative certainty" (Williams 276) suggests that Hume has a tendency to collapse truth into demonstrative certainty. Much to the contrary. In the penultimate paragraph of *Enquiry* 5, Hume writes: "Had not the presence of an object instantly excited the idea of those objects, commonly conjoined with it, all our *knowledge* must have been limited to the narrow sphere of our memory and senses" (EHU 5.2.21, emphases added; SBN 55). Here, "knowledge" implies the "correspondence" Hume is discussing; I do not see how Hume could understand this "correspondence" except as implying truth.

Hume's characterizations of beliefs that result from causal inference often suggest beliefs that are true (SJ 61, 73, 77), even though he does not consider these beliefs demonstratively certain: "No matter of fact can be *proved* but from its cause or its effect. Nothing can be *known* to be the cause of another but by experience" (Abs. 21, emphases added; SBN 654); a person who stops his journey at a riverbank "*foresees* the consequences of his proceeding forward; and his *knowledge* of these consequences is convey'd to him by past experience" (T 1.3.8.13, emphases added; SBN 103). The italicized terms imply success in securing true belief. (Other such passages are cited in the final part of my response.) Hume allows that beliefs about unobserved matters of fact, though far short of demonstratively certain, fasten on

to truth. This is what we should expect, since he allows that inductive inference is truth-conducive.

Perhaps Williams takes *Treatise* 1.3.5 as evidence that Hume distances himself from truth (Williams 276–7, 280). Hume observes that "'twill always be impossible to decide with certainty" about the causes of sensory impressions. He adds: "Nor is such a question any way material to our present purpose. We may draw inferences from the coherence of our perceptions, whether they be true or false" (T 1.3.5.2; SBN 84). The early stage of Hume's explanation of cause and effect reasoning sets the context. Causal inference has three components: an impression of sense or memory (1.3.5), the transition to the idea of the cause or effect (1.3.6), and the conversion of the idea into belief (1.3.7). The first paragraph of 1.3.5 is introductory. The third and later paragraphs focus on memory and lay groundwork for the theory of belief in 1.3.7. Of the "three things to explain," the "*First*, The original impression" (T 1.3.5.1; SBN 84) is the subject of a single paragraph, where Hume announces that the "ultimate cause" of impressions is "perfectly inexplicable" (T 1.3.5.2; SBN 84). He begs off, deferring the problem of the external world until 1.4.2, on the ground that he cannot adjudicate between competing hypotheses about the causes of impressions. Hume's recommendation, in such circumstances, is to stick close to experience—here, the impressions themselves. The fact that we cannot discern the truth at the level of theoretical hypothesis does not dictate general indifference to truth.

I have offered an explanation of why there is more talk of stability than of truth in Hume. Even so, truth does not drop out of the picture. There is the commitment to a correspondence between the train of our beliefs and the course of nature. There are the passages implying success at arriving at truth in the case of particular inductive inferences. There are the numerous veritistic passages to which Schmitt calls attention (especially in his parts 1, 4, and 5). There is Hume's avowal in the "Conclusion": "I am uneasy to think I . . . decide concerning truth and falshood, . . . without knowing upon what principles I proceed" (T 1.4.7.12; SBN 271, cf. 1.4.7.3; 265). Williams is somewhat nonchalant about the way in which truth fits into Hume's picture. I press further in part 4 of my reply to Williams, where I take up his emphasis on long-term consensus.

IV. Reply to Williams

1. The Two-stage Model (Williams, part 5)

I trace the development of Hume's theory of justification through a constructive stage, in part 3 and beyond, and a destructive stage, that culminates in 1.4.7 (SJ § I.4). Claiming that the destructive stage is by and large misguided, I appeal to Hume's perversity in explaining it. Williams thinks this picture mistaken in most every

respect: Hume's epistemological investigations issue in an outcome that is constructive through and through. Williams thus avoids leaning on extra-philosophical considerations. These features of his interpretation have considerable appeal.

Williams's central criticism is this: Hume sees "no need . . . to abandon his inquiries into human nature," his pursuit of the science of man (Williams 268; cf. 283–4); yet "according to Loeb, Hume offers an elaborate naturalistic epistemology with real normative bite, only to take it all away" (Williams 281; cf. 283).[39] Whatever the merits of this objection to my book, Williams's problem vanishes once we attribute to Hume either the average or temporal stability version of a stability theory of justification (part 3 of my reply to Schmitt).[40] Even persons who achieve the intense reflection characteristic of 1.4.7 engage in such hyper-reflection only a small fraction of the time. On the temporal stability version, all belief is destabilized and highly unjustified during intensely reflective periods, but many beliefs are highly stable and highly justified during less reflective periods; loss of justification is transitory or temporary.[41] Consider the average stability version, again applied to an intensely reflective subject. Since intense reflection is rare, many of the subject's beliefs possess sufficient average stability to be justified, and indeed justified to an equal degree, during both intensely reflective and less reflective periods. (Justification is stable, but never reaches so high a degree as during less reflective periods in the temporal stability version.) *On either the average or temporal stability account, sufficient stability and justification survive 1.4.7 to underwrite the enterprise of Books II and III.*

It is the average and temporal stability versions that are best motivated by the value of stability and the ratification argument. They conform to Schmitt's requirement that the justification of a belief held during relatively unreflective periods ought to depend, at least in part, on the tendency of the mechanism that produces it to yield stability during unreflective periods themselves. Schmitt's insightful discussion of stability during these periods dovetails with Williams's insistence that I exaggerate the destructive results of 1.4.7 by overlooking the temporary character of the destabilization of belief (Williams 284, 288–9), but also supplies the alternatives to peak reflection accounts that defeat Williams's objection.

There remains a clear sense in which intense reflection is destructive. In developing this point, I continue to bracket the question (discussed in part 3 below) of whether intense reflection, though destabilizing, might be instrumental for stability. On the temporal stability version, intense reflection destroys all justification during the periods when it holds sway; on the average stability view, the instability under intense reflection lowers the average stability and hence the degree of justification for any belief. *On either the average or temporal stability account, intense reflection takes its toll, detracting from justification, but is not so destructive as to render moral science otiose.* This result preserves the fundamental structure of the two-stage model.

Consider a person who is highly reflective, though not to the point of the hyper-reflection in 1.4.7. On either an average or temporal stability account, this person has many highly justified beliefs during every temporal period. In comparison to such a highly reflective person, the intense reflection in 1.4.7 entails lower overall justification than that of the highly (but not hyper-) reflective person. Absent offsetting stabilizing effects down the line, intense reflection diminishes the degree to which such subjects are justified (in the average stability version) or reduces the amount of time during which such subjects are justified (in the temporal stability version). Either way, the intensely reflective subject is less well off, epistemically, than a subject who attains the highest level of reflection short of destabilizing hyper-reflection.

In my modified interpretation, there is no need to explain why Hume should "push [external world scepticism] to the point at which . . . it undercuts the naturalistic reconstruction of justification" (Williams 283). But Williams could question why there is so much as a give-back. I offer two sets of comments—about the structure of the two-stage development and about the contribution of the appeal to perversity—that bear on this question.

I do allege that Hume relishes paradox and loves to shock—both in his substantive results and in their staging—offering this hypothesis as "ancillary" to the two-stage model (SJ 22). Hume has a constructive and destructive side; he indulges the destructive side, allowing it too much play (cf. SJ 16, 251). Williams, who has little sympathy with assigning an explanatory role to Hume's personality (Williams 267, 281, 283–4), portrays my position in terms that are too stark. I do think Hume involved in infelicities, to be described case by case, that leave his epistemology in a less satisfactory condition than he could have achieved. With one exception (part 3 below), I do not offer features of Hume's intellectual personality as a complete explanation for these missteps. I do suggest that Hume was actively seeking destructive results (SJ 36, 217, 222, 224), but he also locates arguments for them. After identifying an argumentative lapse, I typically provide an explanation of why the line of argument would have appeared plausible to Hume. These explanations appeal to Hume's *philosophical* views, explicit or implicit. Finding some satisfaction in a startling result to which the misstep leads, Hume was not motivated to reexamine his steps (SJ 206–7, 216, 217, 229). I am decidedly not saying that "there is no philosophical explanation" for the misstep or that Hume is "just frivolous" (Williams 283). There is an interplay between perversity and integrity, factors weighted somewhat differently in different cases. If this amounts to "switch[ing] from the integrity to the perversity explanation" (Williams 283), so be it. I offer illustrations in parts 2 and 3 below.

I turn to the structure of the destructive stage. One point concerns its relationship to the constructive stage. Williams describes my account as follows: "The stages are not independent. Rather, the second stage grows out of the first. As Loeb

puts it, Hume's pursuit of his constructive project 'gives rise to a destructive result' (SJ 12)" (Williams 278). I did not intend to put any dialectical weight on the phrase "gives rise to." I write that Hume "comes to conclude" (SJ 15) or "comes to maintain" (SJ 32) that justification cannot be achieved, at least for the reflective person. This phraseology is intended to suggest that the outcome of the destructive stage is not genuinely dependent on the constructive stage. Obviously, since I contend that the destructive stage rests on philosophical mistakes that allow perversity to come into play, I do not think it determined by the constructive stage.

A second point concerns the components or internal structure of the destructive stage itself. I believe my book clear that its main components are the "manifest contradiction" (T 1.4.7.4; SBN 266) and the "very dangerous dilemma" (T 1.4.7.6; SBN 267) in "Conclusion of this book," together with the arguments in 1.4.4 and 1.4.1 (sections to which Hume explicitly refers in 1.4.7) that respectively lay the foundation for these results (SJ 14, 30, 32, 216–7). Williams takes the destructive stage to consist in Hume's "external world skepticism"—the arguments of 1.4.2 and 1.4.4 (thus including the manifest contradiction in 1.4.7). This distorts my position. Section 1.4.2 is essential to the account Williams targets, but inessential to my account of the destructive stage. Section 1.4.1 and the dangerous dilemma are essential to my account of the destructive stage, whereas Williams, in a single paragraph, sloughs off these developments under the rubric of Hume's rejection of "maxims" (Williams 293). I do see 1.4.2 and 1.4.4 as together serving a destructive purpose, putting in question the intelligibility of the subject matter of natural philosophy. There are thus *destructive results* in part 4 that are not proper to the *destructive stage* of the development of Hume's *theory of justification*. On my picture, 1.4.4 contributes to two rather different destructive results: one, via the manifest contradiction, with respect to epistemology; and another, in conjunction with 1.4.2, with respect to natural philosophy. I pursue these matters in the next part.

2. "External World Skepticism" (Williams, parts 5 and 7)

I summarize Williams's characterization of "external world skepticism" as follows (Williams 278, 279–80, 281–2). (1) Hume begins with "naive realism," the vulgar belief in the continued and distinct existence of the immediate objects of awareness; the phenomenon of double vision (T 1.4.2.45; SBN 210–1) shows "the mind-dependence" of these objects and hence the falsehood of this belief. (2) "We are thus led to distinguish perceptions and objects," the hypothesis of double existence (cf. T 1.4.1.46, 52; SBN 211, 215). (3) "Initially, we take perceptions and their objects to be exactly resembling"; this "naive form" of representative or indirect realism is destabilized by appeal to the "observer-dependence of secondary qualities." (Williams

presumably has in view T 1.4.4.3–4.) (4) We are thus "[led] . . . to embrace 'the modern philosophy"; this "sophisticated version of representative realism" is destabilized by appeal to the inconceivability of primary qualities in abstraction from secondary qualities. (Williams has in view T 1.4.4.5–14.) (5) It is this result, coupled with the irresistibility of belief in the external world, that generates "the direct and total opposition" (T 1.4.4.15; SBN 231) between fundamental belief-forming mechanisms at 1.4.4 and thus the manifest contradiction at 1.4.7. On Williams's reading, Hume confronts the vulgar belief and two versions of an hypothesis of double existence: the naive version attributes to objects primary and secondary qualities; the sophisticated version attributes to objects primary qualities alone. These are "philosophical replacements" for the vulgar belief. Further, it is "causal reasoning" that destabilizes both the vulgar belief and the naive form of representative realism, and thus "pushes further along the road to scepticism and the 'manifest contradiction.'"[42]

Williams writes: "According to Loeb, the driving force behind Hume's sceptical turn is the deep conflict that Hume claims to discover between causal reasoning and our disposition to believe in the existence of 'bodies'" (Williams 278). Put aside that this ignores the centrality of 1.4.1 and the dangerous dilemma in my account of the destructive result. Williams presents the manifest contradiction as the culmination of a continuous dialectical development, "external world skepticism," that encompasses all the key arguments in 1.4.2 and 1.4.4. If it is incumbent on me to explain why Hume should allow external world skepticism to tarnish the constructive stage, this sets the explanatory burden high indeed.

I protest. Hume does not portray either the naive or the sophisticated versions of representative realism as outcomes of the falsehood of the vulgar belief in body. At the close of 1.4.2, Hume announces that "going upon the supposition" that "there is both an external and internal world," "I intend to examine some general systems both antient and modern, which have been propos'd of both" (T 1.4.2.57; SBN 218). "Of the antient philosophy" intervenes between sections 1.4.2 and 1.4.4. This seems odd if the two forms of representative realism flow from difficulties for the vulgar belief. In 1.4.4, Hume introduces the modern philosophy by formulating its "fundamental principle" (T 1.4.4.3; SBN 226). He grants that this principle follows from the observer dependence of secondary qualities, without offering any explanation of why the modern philosophy attributes primary qualities to objects in the first place. In addition, 1.4.4 is primarily an attack on realist or materialist conceptions of extended, solid bodies—"philosophical theories of matter" (Williams 283)—whereas the hypothesis of double existence postulates "objects" (T 1.4.2.46,47,52; SBN 211, 212, 215) or "external objects" (T 1.4.2.54; SBN 216) over and above perceptions. The ambiguity is studied: there is belief in "another existence" (T 1.4.2.48; SBN 213), "attributing a continu'd existence to something else, which we call *objects*" (T 1.4.2.52; SBN 215).

What sorts of objects? There are conceptual constraints (SJ 163–4). Witness the penultimate paragraph of 1.4.2: in embracing the hypothesis of double existence, "Philosophers . . . arbitrarily invent a new set of perceptions . . . I say, a new set of perceptions: For . . . 'tis impossible for us distinctly to conceive, objects to be in their nature any thing but exactly the same with perceptions" (T 1.4.2.56; SBN 218). This difficulty traces back to Hume's 1.2.6 argument for the conclusion that "'tis impossible for us so much as to conceive or form an idea of any thing specifically different from ideas and impressions" (T 1.2.6.8; SBN 67–8). In the following paragraph, Hume refers the reader to 1.4.2 for fuller discussion. The second paragraph of 1.4.2 incorporates a note to 1.2.6. Hume picks up the thread late in the section, at 1.4.2.54 (SBN 216) and in the penultimate paragraph. The point is reiterated in 1.4.5 (T 1.4.5.19; SBN 241). Thus, the second set of objects "resembl[e] these perceptions *in their nature*" (T 1.4.2.48, emphasis added; SBN 213). Taking the objects to possess primary qualities is inconsistent with the conceptual strictures in 1.2.6. The sophisticated version of representative realism is conceptually deficient twice over: once, because we cannot conceive of anything other than perceptions; and once again, because we cannot conceive of primary qualities in abstraction from secondary qualities. I agree that Hume is more accommodating to the vulgar belief in body than to philosophical theories of matter (Williams 283).

In 1.4.2, the reflections on double vision that destabilize the vulgar belief lead to the supposition of a second set of *perceptions*. Here it is Williams who does not pay sufficient heed to Hume's conceptual skepticism. "Objects" of any sort other than perceptions are inconceivable. Granted, insofar as it involves a second set of "objects," the hypothesis of double existence is the conceptually legitimate analogue of representative realism; it thus has some claim to be called the "philosophical system." At the same time, its second set of perceptions are not the objects of a realist theory of matter. (At SJ 215–6, I skirt this distinction.) The hypothesis of double existence, in Hume's presentation, preempts the emergence of representative realism; the replacements for the vulgar belief play themselves out before the conceptual difficulties specifically for the notion of matter, as proffered in 1.4.4, can arise. The results in 1.4.2 and those in 1.4.4 are independent dialectical developments. Williams's question for me—that of why Hume rides external world skepticism so hard—divides into parts.

I begin with 1.4.2. It is clear that this section issues in a Pyrrhonian result: "the sceptical doubt arises naturally from a profound and intense reflection. . . . Carelessness and in-attention alone can afford us any remedy" (T 1.4.2.57; SBN 218). To a significant extent, Hume embraces skepticism about the senses as a result of pressures from his own philosophical positions. A number of assumptions constrain 1.4.2. The first is the conceptual assumption we have considered: we cannot conceive of objects specifically different from perceptions. Two assumptions are metaphysical: the immediate or direct objects of awareness are impressions

(SJ 208–9); and impressions do not have a continued and distinct existence (the upshot of reflection on double vision). A fourth is epistemological: the conceptual constraint aside, legitimate causal inference is restricted to induction from observed conjunctions (SJ 193–4)—theoretical inference is precluded (part III). The ontological assumptions show that the vulgar belief is false; the conceptual and epistemological assumptions target double existence. As a consequence of these constraints, both the vulgar belief and the opinion of double existence are unsatisfactory: "'Tis impossible upon any system to defend either our understanding or senses" (T 1.4.2.57; SBN 218). Hume tolerates this result, in large part, because his ontology, his analysis of double vision, his conceptual skepticism, and his strictures on legitimate causal inference, seem to require it.

Williams suggests that my interpretation faces a particular burden: "to explain why Hume [is] so much more receptive to external world scepticism than he is to inductive scepticism" (Williams 281). Although 1.4.2 is not proper to the destructive stage, it does have a similar tenor (SJ 216). Hume's anti-Rationalist argument, in the case of induction, is that there is neither a deductive nor (non-question-begging) inductive route to belief in the unobserved (Williams 274). This does not generate unqualified skepticism (either theoretical or suspensive) about induction, because inductive inference is rooted in custom (Williams 273, 274), which is conducive to stability and thus justifying. The epistemological assumption—precluding theoretical inference—detrimental to the hypothesis of double existence is perhaps similar in spirit to the concern about circularity in 1.3.6.[43] Even so, Hume's assumptions in 1.4.2 go well beyond those required to reach an anti-Rationalist result about the external world. In 1.3.6, there is no analogue to the conceptual assumption in 1.4.2. There is also no analogue to the second metaphysical thesis, that impressions do not have a continued and distinct existence. We hardly find Hume arguing in 1.3.6 that the uniformity of nature is false, or that unobserved causes (effects) of observed effects (causes) do not exist. Perhaps most significantly, though Hume does not argue for the falsehood of the vulgar belief until paragraph 45 of 1.4.2, its falsehood is entrenched in early stage-setting in paragraphs 14 and 15 (SJ 194–7). The cases of induction and the external world are simply different. Convinced that the vulgar belief is false, Hume indeed "allows the principles of natural belief to generate instability" (Williams 281), that is, he allows that the vulgar belief—unlike inductive inference—results from principles that tend to produce instability.

Why is Williams's convinced that I am wrong to present "Hume as having quite different attitudes towards inductive inference and belief in the external world" (Williams 281)? His strategy is to compare 1.4.2 to 1.3.14 rather than to 1.3.6: "Hume thinks that the ordinary conception of causation—which is implicated in our ordinary understanding of the reliability of inductive inference—involves the idea of necessary connection. But the vulgar belief in objective necessary connection is false, or perhaps confused, and fairly readily shown to be so" (Williams

283). Williams grants that "This does not make the cases of induction and naive realism completely parallel. . . . But Hume tries to make the cases appear as close as he can, a point I shall return to" (Williams 283). Williams's idea is that the ordinary conception of causation, as involving a Necessary Connection or Necessity that cannot be reduced to constant conjunction, figures in the ordinary explanation of the reliability of inductive inference. This may be so. But from the perspective of a stability-based interpretation of Hume's epistemology, an approach Williams welcomes, induction is justified because it results from mechanisms that are conducive to stability. Though Hume does believe that induction is reliable (part III), the justification for inductive inference does not rest on the reliability of induction or even on belief in its reliability, much less on an explanation of its reliability.

Hume does not see his argument against Necessary Connection as infecting the positive epistemic status of induction. Inductive inference is "just" at 1.3.6.7 (SBN 89) and 1.3.13.3 (SBN 144). Section 1.3.14 is no cause to flinch; it is directly followed by "Rules by which to judge of causes and effects." The claim that causal inference is just persists well into part 4, as at 1.4.4.1 (SJ 43). The assault on Necessity leaves causation itself similarly untouched. Hume proceeds "to collect all the different parts of this reasoning, and . . . form an exact definition of the relation of cause and effect" (T 1.3.14.30; SBN 169); the next paragraph advances the "two definitions" (SBN 169) of cause. The focus of 1.3.14 is to explain away the belief in Necessity as a confused byproduct of illegitimate imaginative propensities (T 1.3.14.24–29; SBN 167–9).[44] Hume happily "place[s]" (T 2.3.2.4; SBN 409) necessity in constant conjunction; in Hume's view, constant conjunction is necessity enough. We can strip away from the ordinary belief in causation the confused element of Necessity, leaving in place an intelligible account of causation, as constant conjunction. By contrast, paring away from the vulgar belief in body the element that is false—the belief in continued and distinct existence—destroys the belief; no core remains that can amount to a belief in body. Nowhere in part 3 do we find the sort of gloomy mood about either induction or causation that characterizes the final two paragraphs of 1.4.2.

Williams also maintains that 1.4.2 is in an important respect less destructive than I suggest (Williams 283). His point is that we all recur to the vulgar belief outside Pyrrhonian moments; even when destabilized, it exerts a subterranean influence, and so forth (Williams 281; cf. 288). Agreed that the vulgar belief in body has the sort of stability Williams describes. At the same time, consider the identity propensity, the belief-forming mechanism that gives rise to the vulgar belief in cases of the constancy of perceptions. Williams writes: "As Loeb is at pains to point out, this account . . . involves attributing to the human mind susceptibility to all kinds of confusions. Unlike causal reasoning, it doesn't even *look* like the kind of inference that . . . we would intuitively take to be justifying" (Williams 279). Williams does not challenge this assessment. Decidedly, the features of the identity propensity

he enumerates impress Hume as unjustifying. I try to capture this line of thinking, notwithstanding the stability of the belief in body, in terms of the *tendency* of the identity propensity, and that of related imaginative propensities, to generate unstable beliefs (part 4 of my reply to Schmitt); it is these propensities that give rise to the unstable beliefs in material substrata, souls, and the local conjunction of taste with matter (SJ §§ V.3–4).[45] I also provide a rationale for evaluating the epistemic status of beliefs derivatively, in terms of the tendency of belief-forming mechanisms to produce stability (part 4 below).

It is an oversimplification to say that in my interpretation, in light of the confusions inherent in the identity propensity, "the belief in body really isn't justified in *any* sense" (Williams 279). I distinguish between the justifiedness of the belief in body insofar as it arises from constancy via the identity propensity and insofar as it arises from coherence via custom-and-galley (SJ 197–201). The latter mechanism does look like "a kind of reasoning from causation" (T 1.4.2.19; SBN 195), a kind of causal inference as described in part 3. Granted, Hume cannot quite bring himself to say that the belief in body based on custom-and-galley *is* justified. After all, he thinks double vision shows that no perception has continued existence. Yet, he recognizes a disparity between custom-and-galley and the identity propensity. In the first half of the penultimate paragraph of 1.4.2, Hume summarizes his discussion of the causes of the vulgar belief: "They are the coherence and constancy of our perceptions, which produce the opinion of their continu'd existence. . . . The constancy of our perceptions has the most considerable effect, and yet is attended with the greatest difficulties" (T 1.4.2.56; SBN 217). We might put this by saying that the identity propensity, unlike custom-and-galley, is *clearly* unjustifying.

It is the identity propensity that relies on the mind's "susceptibility to all kinds of confusions." I do see Hume as "accepting this feature of the argument" (Williams 279), as investing these confusions with normative import. Hume acknowledges that the belief in body is in a much weaker position—"attended with the greatest difficulties"—insofar as it arises from constancy rather than from coherence. Notwithstanding this asymmetry, the belief is equally irresistible and stable in cases of constancy and coherence. This is awkward, if stability is evaluated in terms of individual beliefs, rather than in terms of tendencies of mechanisms. Williams writes of Hume's "tendency to insist" (Williams 293) that trivial, utterly contingent, properties of the imagination save us from Pyrrhonism. The vulgar belief in cases of constancy rests on a propensity to *mistake* resemblance for identity (SJ 142–4), as Williams agrees. Thus, it is only in cases of constancy that the vulgar belief involves a "fiction" (part 4 of my reply to Schmitt). In the penultimate paragraph of 1.4.2, Hume writes: "'Tis a gross *illusion* to suppose, that our resembling perceptions are numerically the same" (T 1.4.2.56, emphasis added; SBN 217). Hume is far from prepared to shunt aside the sheer confusions internal to the identity propensity as features of "the workings of our own minds" that

"thankfully" stabilize belief (Williams 293). Rather, Hume impugns the identity propensity because the confusions lead to illusions that tend to be unstable.

Hume is in a box with respect to the vulgar belief in body (SJ 196–200). I agree with Schmitt that the falsity of the vulgar belief does not provide any rational support for Hume's seeking the result that the belief is unjustified (Schmitt 316; see SJ 195). But, convinced that the belief is patently false, attributing it to the identity propensity is congenial to Hume. This mechanism, activated by invariable perceptions, does not extend to cases of coherence, which involve change. Hume hits upon custom-and-galley as an explanation of the vulgar belief in these cases. Invoking this mechanism fits nicely with his thesis that all assurance in matters of fact arises from causal inference (SJ 39–44). The upshot is a bipartite explanation of the belief: the identity propensity for constancy, custom-and-galley for coherence. Only the identity propensity is problematic on its face.

Williams would perhaps reject my assumption that Hume's two explanations of the belief in body run on parallel tracks. He maintains that the identity propensity accounts for the concept of body, whereas custom-and-galley "explains our *belief* in body, taking the *concept* of body for granted" (Williams 282). Even if this neat suggestion is correct, it remains that Hume needs to explain how the belief in body arises in the coherence cases. Custom-and-galley is an ingredient in a full explanation of the vulgar belief, whether or not the constancy mechanism has pride of place because it accounts for the concept of body. This suffices to sustain my point about the disparity in Hume's epistemic posture toward the two mechanisms.

Williams's suggestion that the constancy mechanism involves conceptual innovation is nevertheless interesting in its own right. The first part of the four-part explanation of the belief in body in cases of constancy does bear directly on conceptual matters. Hume tells us that the idea of identity seems incompatible with both the ideas of unity and of multiplicity over time (T 1.4.2.26–28; SBN 200). His solution invokes views about the conceptual relationship between change and time or duration (T 1.4.2.29–30; SBN 200–1), views advanced in 1.2.5 (T 1.4.2.29n37; SBN 200–1n1), where Hume works conceptual empiricism hard (T 1.2.5.28–29; SBN 64–5). These conceptual issues are more general than those surrounding the concept of body. The problem is to carve out space for an idea of identity that is engaged in the operation of the propensity to attribute identity to related objects. Williams mentions the beliefs of the ancient philosophers in passing, connecting them to his theme of closing inquiry—they are "exemplary of the ideas that generate interminable clashes of opinion" (Williams 272). He does not mention that Hume attributes these opinions, together with the belief in souls, to the identity propensity (part 4 of my reply to Schmitt). This propensity is the workhorse in part 4, pressed into service to reach negative conclusions about the epistemic status of beliefs in material substrata and souls, as well as the belief in body in cases of constancy.

Granting conceptual innovation in connection with the constancy mechanism, I doubt that Hume's desire to address the conceptual problem "is the key to [his] dissatisfaction with the coherence mechanism" (Williams 282). If the identity propensity is required to supply the concept of body that custom-and-galley takes for granted, it would seem that the belief in body in the coherence cases must inherit the confusion in the operation of the identity propensity, the second part of the four-part system (T 1.4.2.31–36; SBN 201–5). Hume never suggests this; rather, there is the pronounced disparity in his treatment of the two mechanisms. Even in his harshest remarks about custom-and-galley, when Hume pivots from coherence to constancy, he has a light touch: "But whatever force we may ascribe to this principle, I am afraid 'tis too weak to support alone so vast an edifice, as is that of the continu'd existence of all external bodies" (T 1.4.2.23; SBN 198–9). The source of Hume's dissatisfaction with coherence is a large interpretive issue (SJ § VI.3). My own view is that his point is simply that coherence does not apply to "all" instances of the vulgar belief, that it remains to account for cases of constancy (SJ 191–2).

Is Hume right about this? Although the coherence of changing perceptions cannot be a special case of constancy, constancy can be viewed, in Price's phrase, as "monotonous coherence" (SJ § VI.1). This leads to the suggestion that Hume could have supplied a unified treatment of the vulgar belief in body, under custom-and-galley (SJ 212). Not only a unified treatment, but a more constructive one: with its affinity to causal inference, custom-and-galley, unlike the identity propensity, is at least a candidate for a mechanism that confers justification. Had Hume subsumed constancy under coherence, he could have explained all cases of the belief in body in terms of a kindred mechanism of causal inference. The result would have jeopardized his effort to show that the vulgar belief is unjustified (SJ 198–9). I thus agree with Williams, though not on account of the conceptual issues he foregrounds, that Hume would not have welcomed the suggestion that he subsume constancy under custom-and-galley (Williams 283).

In effect, we have traced the Pyrrhonian outcome of 1.4.2 to Hume's metaphysical assumptions. That the immediate objects of awareness are impressions is entrenched doctrine (SJ 208–9). By contrast, the claim that perceptions do not have a continued existence depends upon a discrete argument. This argument is mistaken and Hume could have abandoned it (SJ 209–11). As Williams writes, on my view "Hume's conception of the mind as a bundle of perceptions could and should have led him to take a much more critical attitude towards" the double vision argument (Williams 280). Similarly, Hume could have subsumed constancy under coherence. These two measures would have secured the more constructive outcome in 1.4.2 (SJ 212). (Since the destructive stage, as I conceive it, consists in the manifest contradiction and the dangerous dilemma, it is not my suggestion—contra Williams 280—that the amendments to 1.4.2 would have avoided the destructive results associated with 1.4.7.) Why did Hume not take these steps?

In addressing this question, we revisit the role of perversity in my interpretation. I write: "One would like to identify a philosophical explanation for the total course of Hume's argument" (SJ 193). I do not suggest that Hume accepted the double vision argument simply because he sought a destructive outcome. In so much as considering the possibility that impressions continue to exist when not perceived, Hume was breaking entirely new ground; he was entertaining a radical departure from the theory of ideas, anticipating neutral monist theories that gained some currency only in the late nineteenth century (SJ 210–2). The double vision argument, though mistaken, is seductive; Hume's accepting it is far from an egregious mistake. As far as I know, no one nailed down the error in the argument until Price—in 1940, two centuries after publication of the *Treatise* (SJ 211n41).

Similarly, I do not suggest that Hume failed to subsume constancy under coherence simply in order to vindicate his philosophical convictions by infecting the vulgar belief with the confusion and "difficulties" inherent in the identity propensity. I argue that it never occurred to Hume that the stages of an unchanging object are *causally* dependent upon one another and hence phenomena eligible for causal inference or the allied mechanism of custom-and-galley (SJ § VI.5). I identify both historical and philosophical reasons for this lapse (SJ 203–4), thus explaining why Hume did not subsume constancy under coherence. Again, Hume has not here made some egregious error. In sum, the failure to arrive at the more constructive outcome is rooted in Hume's explicit (in the case of the double vision argument) or implicit (in the case of persistence as a non-causal process) philosophical convictions. Insofar as Hume was enamored of destructive, paradoxical conclusions, he had little motivation to retrace his steps. Perversity does not so much explain how Hume arrived at the destructive results as why he left them standing.

There is a further element in "external world skepticism": the attack on philosophical theories of matter in 1.4.4 (cf. Williams 281–2, 283). The argument proceeds in two steps: first, secondary qualities are nothing but impressions of the mind that do not resemble qualities of objects (T 1.4.4.3–4); second, it is inconceivable that primary qualities exist separate from secondary qualities (T 1.4.4.7–14); thus, it is inconceivable that primary qualities exist in objects (T 1.4.4.6). The first step was common property—in Descartes, Locke, and Berkeley. The second step depends upon Berkeley's position on abstraction (SJ 246). Hume follows the two-step argument through to its conclusion: our conception of matter is deprived of its essential content.

Although 1.4.4 is not a dialectical result of 1.4.2, Hume relies on the combined forces of 1.4.2 and 1.4.4 to conclude that the very subject matter of natural philosophy is problematic. Courtesy of 1.4.2, neither the vulgar belief in body nor the hypothesis of double existence is satisfactory; courtesy of 1.4.4, no form of realism or materialism is satisfactory (cf. SJ 215–6). Hume introduces this theme in the first sentence of 1.4.5: "HAVING found such contradictions and difficulties *in every*

system concerning external objects, and *in the idea of matter . . .* " (emphases added; SBN 232). The italicized phrases refer to the results of 1.4.2 and 1.4.4, respectively. "Bodies"—objects with a continued and independent existence (T 1.4.2.2; SBN 187–8)—can be identified neither with perceptions nor with matter. There is no satisfactory ontology for natural philosophy. This is a result Hume wants to exploit. The opening paragraph of 1.4.5 continues: "The intellectual world, tho' involv'd in infinite obscurities, is not perplex'd with any such contradictions, as those we have discover'd in the natural" (SBN 232). He returns to the theme in Book 2 (T 2.2.6.2; SBN 366). Hume relies on 1.4.2 and 1.4.4 to build a case that moral, if not natural, science, has a clear and comprehensible subject matter—or so he hopes.[46]

Hume welcomes this less than constructive result for physics. The evidence is in his staging. In the Introduction to the *Treatise*, there is no intimation of an asymmetry with respect to the relative intelligibility of the subject matters of natural and moral philosophy. To the contrary, "the essence of the mind [is] equally unknown to us with that of external bodies" (T Intro.8; SBN xvii). Of course, the Introduction assigns the moral sciences pride of place in one respect: "Even *Mathematics*, *Natural Philosophy*, and *Natural Religion*, are in some measure dependent on the science of MAN" (T Intro.4; SBN xv). The natural sciences have the advantage of the possibility of controlled manipulation (T Intro.10; SBN xviii–xvix). This perhaps explains why "the application of experimental philosophy to moral subjects . . . come[s] after that to natural at the distance of above a whole century" (T Intro.7; SBN xvi). It is here that Hume mentions Bacon: "there was about the same interval betwixt the origins of these sciences; and that reckoning from THALES to SOCRATES, the space of time is nearly equal to that betwixt my LORD BACON and some late philosophers in *England*, who have begun to put the science of man on a new footing" (T Intro.7; SBN xvi–xvii). This is where matters stand in the Introduction.

As part 4 unfolds, distinctive perplexities are seen to riddle the domain of natural philosophy. This is a turnabout from the theme of parity rooted in the Lockian claim that the essences of body and mind are equally unknowable. Hume locates a new ground (beyond the dependence of all sciences upon the science of man) for the special standing of the moral sciences. This is a shot across Bacon's bow. He is "the father of experimental physicks" (Abs. 2; SBN 646), whereas Hume's true heroes—Locke, Shaftesbury, Mandeville, Hutcheson, and Butler—pursue the moral science Hume privileges. I do think Hume must have found some pleasure in withholding these perplexities in the "Introduction" and in arriving at them in the first place. According to Williams, neither an explanation of Hume's skeptical moments that appeals to integrity nor one that appeals to perversity "gives Hume much credit for common sense" (Williams 281). I am unsure how to reconcile common sense and Hume's theoretical indictment of the underpinnings of natural science. We have seen that a number of Hume's fundamental convictions exert pressure in the direction of a destructive result in 1.4.2. We can add that his pleasure in the more

general destructive outcome, based on 1.4.2 and 1.4.4, would have contributed to his unwillingness to reexamine his assumptions in these sections.

3. Treatise *1.4.7* *(Williams, parts 6–7)*

The manifest contradiction and the dangerous dilemma are the main components of the destructive stage of Hume's development of his epistemological theory. The context is revealing. Hume takes the paragraph that intervenes between these episodes as occasion seemingly to lament that there is no ultimate "connexion, tie, or energy" (T 1.4.7.5; SBN 266), even though the concept of Necessity is an "illusion" (T 1.4.7.6; SBN 267) and incoherent (part 2 of my reply to Williams). Similarly, the paragraph preceding the manifest contradiction can seem to lament that all belief is founded on vivacity, "which seemingly is so trivial" (T 1.4.7.3; SBN 265). Hume's exposition of his theories of causation and belief in part 3 hardly prepares the reader for the tone of these passages. Also, I argue that Hume's implication of causal inference in the 1.4.4 argument about secondary qualities is artificial; he there strains to impugn his own favored belief-forming mechanism (SJ § VII.2). (This is the one juncture where Hume's love of paradox determines the course of his argument.) All in all, Hume does overreach in his build-up to the dilemma.

In downpeddling the role of the dangerous dilemma, Williams relieves himself of the urgency of explaining how Hume came to embrace it. The dilemma draws essentially on the claim that all probability reduces to zero (T 1.4.7.7 and n53; SBN 267–8). Even though he builds its engine in 1.4.1, Hume does not disclose the dilemma until 1.4.7. This is reminiscent of the presentation of his jolt to physics—no hint in the Introduction of the results to come in 1.4.2 and 1.4.4 (part 2 of my reply to Williams). Hume's argument for the reduction of probability is famously problematic (part 1 of my reply to Schmitt). Even here, I do not think Hume "just frivolous." In Hume's theory of belief, vivacity—though a single variable that only admits of differences in degree—is the bearer both of propositional attitudes and features of propositional content. Hume's theory of probability is a special case, in which vivacity is identified with both degree of confidence and estimate of statistical likelihood. This fosters a mistake; the argument for the reduction of probability to zero trades on the two roles (SJ 228–9). Again, a line of argument that would have appeared plausible to Hume leads to a negative result that he finds congenial; in these circumstances, there is little self-monitoring, little effort to uncover arguments that go amiss.

Williams treats the dangerous dilemma as a development apart: "[L]et me note a feature of Hume's 'sceptical' conclusion that is too little remarked on: his insistence on the impossibility of guiding our inquiries by rules or maxims" (Williams 293).

Williams's emphasis on maxims is odd. Part 3 is replete with rules: in the discussions of the relation of resemblance (T 1.3.9.6; SBN 110, cf. 1.3.10.11; 631–2), the probability of causes (T 1.3.12.24; SBN 141–2), the fourth kind of unphilosophical probability (T 1.3.13.7–12; SBN 146–50), and rules by which to judge causes and effects (T 1.3.15.1; SBN 173–5). Hume is happy to apply the term 'maxim' in a variety of contexts. I doubt that the language of "maxim" and "rule" (T 1.4.7.7; SBN 268) is reason to downplay the dilemma's intended force. The dramatic closing pages of the "Conclusion," which begin just following Hume's statement of the dilemma, seem an extreme response to a rejection of maxims.

The dangerous dilemma is an extension of Hume's line of thought in the manifest contradiction. Section 1.4.4 contributes to two destructive outcomes: joining forces with 1.4.2, it casts a shadow within ontology, on the subject matter of natural philosophy (part 2 of my reply to Williams); it also sets the stage for the dilemma, which undermines Hume's favored epistemological distinctions and hence scientific methodology. Hume finds that "two operations" of the imagination lead to contradiction, though "equally natural and necessary" (T 1.4.7.4; SBN 265). He writes in the first sentence of the following paragraph: "This contradiction wou'd be more excusable, were it compensated by any degree of solidity and satisfaction in the other parts of our reasoning"; unfortunately, "the case is quite contrary" (T 1.4.7.5; SBN 266). The manifest contradiction, albeit involving fundamental beliefs, might itself be "excusable" because it is localized, restricted to assumptions germane to natural philosophy (SJ 30).[47]

It is significant that the manifest contradiction is deployed between paragraphs in which Hume presses topic-neutral considerations. The professed concern that belief is founded in vivacity applies to all belief, irrespective of subject matter. The related worry about "the imagination, or the vivacity of our ideas" (T 1.4.7.3; SBN 265) introduces the manifest contradiction. The professed metaphysical concern—that our wish to "arrive at . . . that tie, which connects" cause and effect is frustrated (T 1.4.7.5; SBN 266)—applies to natural philosophy and moral philosophy alike. This thought introduces the dangerous dilemma, where the question is a general one: "how far we ought to yield to . . . illusions" of the imagination (T 1.4.7.6; SBN 267). The dilemma is systemic, applying to beliefs irrespective of their content or subject matter. The manifest contradiction raises questions Hume cannot answer: "How then shall we adjust those principles together?" (T 1.4.7.4; SBN 266). We find the same note toward the close of the dilemma: "We have, therefore, no choice left but betwixt a false reason and none at all. For my part, I know not what ought to be done" (T 1.4.7.7; SBN 268). In presenting the manifest contradiction as nothing but the final outcome of "external world skepticism," Williams obscures its role as a way-station to the dilemma.

Hume describes the horns of the dangerous dilemma as either relying on the whole of the imagination (T 1.4.7.3,6; SBN 265, 267) or "adher[ing] to the

understanding, that is, to the general and the more establish'd properties of the imagination" alone (T 1.4.7.7; SBN 267). His terminology recapitulates that at SBN 117n1 (T 1.3.9n22) and 371n1 (SJ 53–5). It also harkens back to the first paragraph of 1.4.4; the "permanent, irresistible, and universal principles" presumably constitute "the general and more establish'd properties of the imagination." Williams agrees that 1.4.4.1 is a principal locus of Hume's constructive attempt to draw epistemic distinctions (Williams 276–7, 285–6). The dilemma is a further development in this enterprise, but one that reverses course. It is because 1.4.1 targets the "understanding" (T 1.4.1.1,5,12; SBN 180, 182, 187) that it can feed the "false reason" (T 1.4.7.7; SBN 268) horn of the dilemma (SJ 85–6, 223). The dilemma does not change the subject, to maxims; it is an assault on systematization of normative epistemic distinctions. (This is not a failure to identify Rationalist "foundations" for our epistemic commitments; it is a failure to so much as systematize those commitments from a naturalistic perspective.) It is not surprising that Hume is "ready to reject all belief and reasoning, and can look upon no opinion even as more probable or likely than another" (T 1.4.7.8; SBN 268–9). The enterprise epitomized at the beginning of 1.4.4 has run aground.

Williams will have none of this; in 1.4.7, Hume reaches the sweet spot, a Pyrrhonian crisis whose natural result is to temper credulity and dogmatism—a positive, because stabilizing, development. Williams advances a powerful account of the "Conclusion." He beautifully weaves together Humean themes: the temporary character of Pyrrhonian skepticism, trivial properties of the human mind, and self-correcting features of the cognitive faculties. Much as in his account of "external world skepticism," Williams invites us to find a rich, directed philosophical structure. I cannot here do Williams's position justice, but I do not think Hume this tidy.

Williams agrees Pyrrhonian developments are radically destabilizing during intense reflection. Do they also promote stability that at least offsets these effects? Here I take up the issue, bracketed to this point, of the instrumental value of intense reflection. Once we adopt an interpretation in terms of average or temporal stability (part 3 of my reply to Schmitt and part 1 of my reply to Williams), there is sufficient epistemic justification to sustain the science of man even if Pyrrhonian reflection is not instrumental to stability; such an instrumental effect would nevertheless be a nice bonus.

Much of Hume's initial maneuvering is curious on Williams's interpretation. Hume would "be a loser in point of pleasure" (T 1.4.7.12; SBN 271) were he not to indulge his curiosity and pursue his ambition. Superstition would lead Hume beyond common life, even if philosophy did not do so. Hume recommends philosophy for a "guide" because it is "safest"—"the errors in religion are dangerous; those in philosophy only ridiculous" (T 1.4.7.13; T 272)—"and most agreeable" (T 1.4.7.13; SBN 271). A defense of philosophy along these lines would seem superflu-

ous if the intense reflection in the dangerous dilemma leads to higher and more expansive plateaus of stability.

Let me turn to the specific theme that Hume "seems positively to *welcome* [Pyrrhonism]" because it "is the corrective for our dispositions toward credulity and dogmatism" (Williams 290), dispositions that manifest themselves within philosophy itself. I begin with the intriguing claim that Hume pursues a theme of "self-correction through conflict" in both 1.3.13 and 1.4.7 (Williams 291–2). It is here that Williams returns to his suggestion (Williams 283–4) that Hume presses a parallelism between inductive inference and the vulgar belief in body, in that serious conceptual issues impact both (part 2 of my reply to Williams). It is integral to my own perspective that conflict is a source of uneasiness and provides an impetus for adjustment. Hume presents the manifest contradiction as an instance where we seem doomed to oscillation (SJ 8–9). In other instances, such as the probability of causes, we manage to extract a single judgment from opposing beliefs. In 1.3.13, "general rules" can rescue us from rash generalizations based on accidental (coincidental) circumstances. This is the case Williams has in view. The adjustment, however, consists in *expanding* the scope of induction to include higher-order inductive inferences about the success or failure of classes of first-order inductions (Williams 273; SJ § IV.2). The trick is to make more use of induction—by making more general use of induction—not less. In the case of Pyrrhonian skepticism, the effect of any "self-correction" is to chasten intense reflection, thereby *containing* Pyrrhonism itself. This does not show that Williams is mistaken in seeing Pyrrhonism as a corrective; it does suggest a sense in which part 3 is more kind to inductive inference than is part 4 to the intense reflection characteristic of external world skepticism and the dangerous dilemma.

Williams writes that Hume makes the point about the corrective power of Pyrrhonism "even more clearly" (Williams 290) in the first *Enquiry* than in the *Treatise*. He is not the first to fall back on the *Enquiry* to make a case for the chastening effects of Pyrrhonism.[48] Even in this work, Hume advances the case for the utility of Pyrrhonism somewhat speculatively: "There is, indeed, a more mitigated scepticism or ACADEMICAL philosophy, which may be both durable and useful, and which *may, in part*, be the result of this PYRRHONISM" (EHU 12.3.24, emphases added and deleted; SBN 161, cf. 12.3.25; 162). Care is needed not to read the *Enquiry* passages into the *Treatise*.

Granted, Hume expresses the "wish [that] we cou'd communicate to our founders of systems, a share of [the honest gentlemen's] gross earthy mixture" (T 1.4.7.14; SBN 272). I do not see a textual basis in the *Treatise* for Williams's claim that "An experience of the force of profound sceptical reflections is *the means* by which the 'fiery' tempered philosopher can acquire a touch of the country gentleman's earthiness" (Williams 290, emphasis added). Later in the paragraph, we have Hume's discussion of a "true sceptic" who "finds himself in [an] easy disposition," "studies

philosophy in this careless manner," and is thus "diffident of his philosophical doubts" (T 1.4.7.14; SBN 273). Even if we suppose that "true" skepticism tempers speculation and dogmatism, there is no suggestion that true skepticism results from Pyrrhonism, and thus no suggestion that Pyrrhonism is "an integral component" (Williams 268) of Hume's constructive epistemological project. Rather than claim tempering effects as a virtue even of true skepticism, Hume returns to the theme of pursuing pleasure: the true sceptic "will never refuse any innocent satisfaction" (T 1.4.7.14; SBN 273). This seems far removed from Williams's account of why Hume "feel[s] entitled to proceed with his science of man" (Williams 284).[49]

It might be urged that what matters is the trajectory of Hume's thinking; that on the topics at hand the *Treatise* takes steps on the road to the first *Enquiry*, where we find Hume's best philosophical thought. Perhaps, but I am unsure that the environment of the *Treatise* is hospitable to the suggestions in *Enquiry* 12.3. Recall Hume's exploitation of the "contradictions and difficulties in every system concerning external objects, and in the idea of matter" (part 2 of my reply to Williams). Hume does not suggest that the reflections that reveal obscurities in natural philosophy are virtuous, giving rise to a healthy chastening—tempering dogmatism and excess in hypotheses; nor that moral philosophers are disadvantaged for want of contradictions in the subject matter of their inquiries. Rather, the Pyrrhonian reflections in 1.4.2 and 1.4.4 leave natural philosophy worse off; any benefits accrue to the relative standing of moral philosophy. In much the same way that 1.4.2 and 1.4.4 detract from natural philosophy, 1.4.7 is a blow to a normative systematization of belief-forming mechanisms, to methodology.

This specter vanishes in the first *Enquiry*. Hume mentions the "natural weakness of human understanding" (EHU 12.2.21; SBN 158), but there is no hint of the reiterated corrections for mistakes in estimates of probabilities that drive the reduction of probability to zero. As a consequence, the dangerous dilemma does not appear in the *Enquiry*. This leaves space for the thought that "philosophical decisions are nothing but the reflections of common life, methodized and corrected" (EHU 12.3.25; SBN 162). In the *Treatise*, the possibility of such a systematization is lost to the dilemma. Hume is left to contribute to moral philosophy "in some particulars" and "in *particular points*" (T 1.4.7.14,15; SBN 273). The codification of a method cannot be achieved. In my book, I seek to interpret the *Treatise* on its own terms (SJ ix).

The entire project of drawing epistemic distinctions takes a different shape in the first *Enquiry*. In the *Treatise*, the aim is to partition associative mechanisms into two camps: understanding or reason—the home of causal inference—and mere propensities of the imagination (SJ § II.3). Part 4 of the *Treatise*, which is devoted to these imaginative propensities (SJ 246), is reduced to a single section in the *Enquiry* (EHU 12); the identity propensity, the chief imaginative propensity in the *Treatise* (part 2 of my reply to Williams), drops out of sight. By contrast, the

primary epistemic focus of the *Enquiry* is to draw distinctions within reason or the understanding itself, that is, to differentiate legitimate and illegitimate forms of causal or inductive inference; these distinctions power the arguments about miracles and the religious hypothesis. This shunts aside the broader objective of drawing epistemic distinctions between the understanding and the imagination, the project that runs afoul of the dangerous dilemma. To clear the way for portraying Pyrrhonism in a more constructive light, Hume abandoned the dilemma. This is a measure of how hard Hume does want to ride its destructive implications in the *Treatise*.

The Berkelian argument against matter, together with highly abbreviated elements of 1.4.2, remain in the first *Enquiry* (EHU 12.1.6–17; SBN 151–5). They are not, however, utilized to establish an asymmetry in the standing of natural and moral philosophy. In the *Enquiry*, the various skeptical worries about the external world in 12.1 are thereby free to serve as representatives of Pyrrhonism. These forms of skepticism are not slight, but, unlike the dangerous dilemma, they do not systematically strike at the heart of the *Treatise's* project of distinguishing the understanding and the imagination epistemically.

In addition to the reduction of probability to zero and the dangerous dilemma, many *Treatise* positions that run off the rails and could stand correction are dropped in the first *Enquiry*. Although the Berkelian argument against matter remains, the innovation to implicate causal reasoning in the argument does not (SJ 220). The possibility that impressions have a continued and distinct existence (and hence the vulgar belief in body in the form it takes in the *Treatise*) receives no play. A linchpin of 1.4.2—double vision—is reduced to one of "the more trite topics, employed by the skeptics in all ages" (EHU 12.1.6; SBN 151). Hume appeals to a single, primitive instinct to explain the belief in body (EHU 12.1.7; SBN 151); there is no mention of coherence and constancy and no need for the identity propensity and a host of associated difficulties (SJ 172–3, 212–3). Williams speculates that Hume came to recognize that he pressed too hard for parallelism between inductive inference and the belief in body, leading him to dilute or omit these discussions (Williams 296n18). This strikes me as unlikely, both in light of the disanalogies in the two developments (part 2 of my reply to Williams) and the thin textual evidence for Pyrrhonian self-correction in the *Treatise*.

My alternative suggestion is that the more mature Hume sensed that the *Treatise* discussions eviscerated in the first *Enquiry* were substantive vulnerabilities in the "juvenile work" (EHU Adv; SBN, 2). The 1751 letter to Gilbert Elliot of Minot (L 1:158) encourages the thought that Hume was unhappy with the intricacy of the *Treatise* and that the *Enquiry* differs principally "in the expression" (EHU Adv; SBN, 2). This is an extra-philosophical explanation of the differences between these works. Consider, however, the intricacies that Hume jettisons in the *Enquiry*: the reduction of probability to zero, any role for the identity propensity, the divi-

sion of labor between coherence and constancy, the double vision argument, the assimilation of the argument against matter to causal inference, the dangerous dilemma. These developments contain the very missteps that lead to unnecessarily destructive results. Perhaps Hume recognizes as much: "I was carry'd away by the Heat of Youth & Invention" (L 1:158). I do not think it too much to suggest that the youthful Hume enjoyed parading novel difficulties and thus did not subject his arguments to the fullest scrutiny. In later years, Hume appreciated that the arguments might well be deficient.

4. Long-term Consensus (Williams, part 6)

There are two planks in Williams's attack on my claim that Hume arrives at a destructive epistemological outcome. I have considered the contention that I overestimate the extent to which the outcome is destabilizing. In my book, I exaggerate the skeptical results simply in virtue of relying on peak reflection versions of a stability theory. I should have turned to average and temporal stability versions (part 3 of my reply to Schmitt and part 1 of my reply to Williams). This response undercuts the motivation for the second plank, the contention that I am blind to the kind of stability Hume has in view (Williams 284–8). Williams writes that an "overly simple notion of doxastic stability . . . leads [Loeb] to an overestimation of the project's sceptical results" (Williams 288; cf. 286). Williams's diagnosis of the problem is that I focus on individual stability rather than consensus in the (very) long-term. My overestimation of the skeptical results, however, is fully explained in terms of my invoking an overly simple notion of *individual* stability. The real culprit is my adoption of peak reflection accounts; the average and temporal stability accounts that repair the damage are theories of individual justification.

Similarly, even if Pyrrhonism serves a corrective function (see the previous part), it does not follow that Hume identifies the goal of inquiry with stability in the long term. Those who experience Pyrrhonian doubt presumably benefit from any chastening effects, so that Pyrrhonism is self-correcting for the individual, whether or not long-term social consensus is the goal of inquiry. Williams's emphasis on the (very) long term is critical to his strategy for mitigating the import of the periodic bouts of instability, mere specks over many thousands of years. As we have seen, the average and temporal stability theories of individual stability preserve justification without any appeal to the long-run.

What remains of Williams's manifesto for long-term consensus? Although some of his evidence is a stretch, there is no need to challenge his insistence that "Hume's interest in consensus can be spotted throughout the *Treatise*" (Williams 31).[50] The question is whether that interest determines his epistemology, whether long-term social consensus is "*the* goal" (Williams 277, 293, emphasis added) that

serves as the basis for Hume's epistemic distinctions. In my view, it is wrong-headed to convert Hume's interest in the long-term into a theory of justification.

My entrée into this issue is by way of the earlier observation that Hume takes causal inference to constitute an *epistemic* achievement—hence the language of *proof*, *foresight*, *knowledge*. We not only land on truth, but fasten on it or secure it. There are many such passages beyond those cited in my part III: causal inference "discover[s]" (T 1.3.2.2; SBN 73), "informs us of" (T 1.3.2.3; SBN 74), and "brings us acquainted with" (T 1.3.9.4; SBN 108) unobserved objects. Also, causal inference is justified, "just" (T 1.3.6.7, 1.3.13.3; SBN 89, 144); to engage in causal inference is to reason "justly" (T 1.4.4.1; SBN 225, EHU 10.1.3; SBN 110—cf. 10.1.10; 113). In all these contexts, Hume has in view the epistemic status of ordinary persons engaged in mundane causal inferences, epistemic successes in the here and now. In terms of the figure of *Enquiry* 5, the pre-established harmony obtains for common persons and in the (very) short term. Indeed, it is Hume's focus on ordinary persons that gives the ratification argument its power (part 2 of my reply to Schmitt).

Even the country gentlemen (Williams 286–7), though not especially reflective, are to be credited with many cognitive achievements. Others do much better, but this is because there are degrees of epistemic success. Williams writes: "If Loeb is to be believed, Hume takes the most (perhaps the only) successful believers to be the stolid and incurious" (286). This is far from the case, since more reflective persons—relying, for example, on "general rules" to root out prejudice—will tend to have more stable beliefs, provided they do not cross the line to hyper-reflection. (See below.)

Williams is dismissive of the idea that Hume seeks to explain epistemic success in terms of individual stability; this strategy is hopeless, "because too undemanding" (Williams 284). I locate a number of Humean resources for blunting this objection (SJ 93–7). Since Williams does not engage these elements of my interpretation, I confine my discussion to an overview of the underlying idea. An objective of individual stability will look too undemanding if we focus on the degree of stability in belief an individual in fact achieves, *de facto* stability. This is the conception in Williams's sights: "For Loeb . . . any individual who can stabilize his beliefs, by whatever means, is capable of justified belief" (Williams 284–5). A prominent theme of my book is that justification is assessed in terms of the *tendency* of belief-forming mechanisms to produce stability (SJ 12–13, 32–3, 88–9). Hume can distinguish between a mechanism's producing stable beliefs in one's actual circumstances and its tendency to produce stable beliefs across a range of relevant circumstances. Williams himself writes of "stability-*conducive* principles" (Williams 293, emphasis added). Such notions as "tendency" and "relevant circumstances" are vague, so that Hume's approach is consistent with a broad class of theories. Hume certainly has no disposition to introduce refinements in the manner of analytic philosophy in recent decades.

The philosophical *strategy* that underpins the class of theories Hume has in view, however, is central to his thinking—both in epistemology and moral philosophy. We have seen, in my response to Kelly, Hume allow that "virtue in rags is still virtue" (T 3.3.1.19; SBN 584). Why is that? Virtue does not benefit society in the circumstances that one is in rags, but it has a tendency (T 3.3.1.10,19,22,23; SBN 576, 577, 578, 579, 584, 586) to benefit society in a range of relevant circumstances. Again, Hume does not spell out details. But much as we evaluate character traits in terms of their tendency to benefit society, we evaluate belief-forming mechanisms in terms of their tendency to produce stable beliefs. The normative implications, which depend upon how the theory is elaborated, need not be nearly as permissive as Williams suggests. I think it premature to conclude that an interpretation in terms of individual stability "is clearly of no use to Hume" (Williams 285).

There is an ambiguity in Williams's round rejection of my interpretation. Is his the (relatively) modest position that, for Hume, individual stability cannot account for epistemic distinctions that apply *to members of the community of inquirers*; or is it the ambitious position that individual stability cannot account even for epistemic distinctions that apply *to common persons*? Williams writes: "Loeb's overly individualistic notion of stability leads him to an interpretation . . . that strikes me as extremely dubious" (Williams 286). This perhaps suggests that Hume has in view an epistemic goal of long-term consensus *rather than* individual stability. Similarly, the country gentlemen "have *nothing to do* with [Hume's] project of stabilizing belief within the community of inquirers, to which they do not belong" (Williams 286, emphasis added). This suggests that a goal of stability for individuals at large, including common persons, is irrelevant to Hume's epistemology. Further, Williams formulates the goal of long-term consensus as applying specifically to "inquirers" (Williams 284–5). A goal of long-term consensus perhaps entails individual stability for those who are parties to the consensus; it does not entail individual stability for common persons, or even for members of the community of inquiry during periods that precede a final consensus. If we read Williams as promoting the stronger position, the goal of long-term consensus within the community of inquirers floats free from any general goal of individual stability. There is a single, and hence autonomous, epistemic goal, long-term consensus within the community of inquirers. I do not know whether this is Williams's considered position, but it is important to be clear about the problems it encounters. There are three large difficulties.

One difficulty is that a free-standing goal of long-term consensus falls short of delivering the normative results Hume wants. Whereas Williams alleges that an epistemological objective of individual stability has implications that are too *undemanding*, an objective of long-term consensus—if detached from considerations of individual stability—threatens to be *indeterminate* in its implications. Which methods of belief-formation are likely to yield, or increase the likelihood

of yielding, long-term consensus? This seems an open question. We have some grip on how to assess tendencies toward stability for individuals, in terms of likelihoods given "relevant" circumstances defined across large populations or sets of circumstances. It is harder to assess the likelihood that a particular set of methods will lead to long-term consensus; the march to posterity unfolds only once.

Perhaps what matters is that if consensus is reached, the goal of inquiry licenses the methods that in fact produce it. Unfortunately, the implications might be unwelcome. In observing that Pyrrhonian doubt may have consensus for a "natural result," Williams writes: "I take [Hume's] talk of 'natural result' seriously" (Williams 291). If a few thousand years of authoritarianism would stabilize belief across individuals, a free-standing goal of long-term consensus would seem to legitimize authoritarian methods of control. Even if long-term consensus "Naturally, . . . must depend on principles that are 'permanent and universal'" (Williams 286), "education" or indoctrination is itself a form of custom or habit operating through repetition (T 1.3.9.16–19; SBN 115–7) and hence founded in the permanent, irresistible, and universal principles. If consensus in the long-term is a sovereign epistemic goal, it is not clear how to prevent an epistemic preference for consensus "by whatever means." Systematic, technologically sophisticated indoctrination might carry the day. It is an exclusive emphasis on (very) long-term consensus that exerts pressure in the direction of evaluating methods in terms of *de facto* stability (albeit at the latest posterity), by precluding consideration of the tendencies of mechanisms to produce stability.

Perhaps the thought is that a goal of long-term consensus at least constrains epistemically respectable methods. In this vein, Williams writes that long-term consensus "will be possible only if there are ways of forming beliefs that can survive sustained critical scrutiny" (Williams 285; cf. 286, 287). But it is *possible* that "long interruptions" (T 1.4.7.14; SBN 273) in inquiry and scrutiny will recur, persist, and eventuate in an uncritical consensus. The likelihood of this outcome seems quite sensitive to "initial conditions," the idiosyncratic particulars of social and historical conditions, both with respect to the belief-forming methods in use and the ways in which those methods and the beliefs that result from them are transmitted. If long-term consensus is the sole epistemological objective, Hume would need to address these dimensions of his position.

There is a second difficulty for the thought that Hume promotes long-term consensus as a free-standing epistemic goal. Hume's arguments to establish the value of stability are squarely rooted in facts about the common individual. The arguments do not appeal to special features of a reformed science pursued in the provinces of the learned. For example, Hume's account of justification focuses on stability *because* steadiness is the natural function of belief (part 1 of my reply to Schmitt); the point of a distinction between justified and unjustified belief in terms of stability is to call attention to circumstances that interfere with beliefs serving

their natural function (SJ § III.4). Moreover, in the opening paragraphs of 1.3.10, this natural function applies to common individuals; steadiness is in the interest of guiding individual action to secure goods and avoid evils. Similarly, some belief-forming mechanisms interfere with the healthy functioning that brings stability (compare part 3); they are likened to "a malady" and hence "contrary to health" (T 1.4.4.1; SBN 225–226, cf. 1.3.16.2; 176). These notions have their primary application to individual organisms, ones that need not possess any special cognitive standing or ambitions.

Hume has a second argument for the value of stability: epistemic obligation is naturalized as deriving from the motivational force of the discomfort in instability (part 3). This consideration is also completely general, applying to common individuals. Were long-term consensus within the community of inquirers the sole epistemic goal, this argument for the value of stability for common individuals would collapse; a common person will not be moved or motivated by the prospect of instability and uneasiness besetting inquirers within a Republic of Letters of which he is not a member. The vulgar are unlikely so much as to consider this prospect. Neither the argument based on the natural function of belief, nor that based on the natural motive to relieve uneasiness, provides any direct motivation for the value of consensus in the long term.

In passages quoted above, Williams insinuates that long-term consensus is the sole epistemic goal at the table. Yet, Williams endorses the second argument for the value of stability: Hume's "investing stability in our beliefs with normative significance" is possible because "he thinks that stability is intrinsically desirable. Putting things the other way around, instability—the presence of conflict and contradiction—produces uneasiness in the mind, and is therefore something that we have a natural interest in avoiding or alleviating" (Williams 277). Williams does not take note of the gulf between this argument, which motivates the value of individual stability even for common persons, and his interpretation emphasizing stability in the form of long-term consensus. Were we to put to the side the evidence of individual stability as an epistemic goal, Williams could not help himself to the two lines of argument for the value of stability I have identified. Would sufficient textual evidence remain to nail down an interpretation in terms of stability of any sort?

The main line of argument in my book extracts a stability interpretation from the "tandem passages." Quite independently of his theory of justification, we have grounds for taking Hume to hold that steadiness is essential to belief: his discussions of firm, solid, and settled belief; of infixing by repetition; and of counterfeit belief (SJ § III.2–3). In the first three sets of tandem passages, as enumerated in part 2 of my reply to Schmitt, Hume's claim that association by the relation of cause and effect produces belief is intertwined—though without his remarking on this fact—with the claim that belief based on causal inference is justified. (Many of

Hume's remarks about the epistemic success of everyday causal inference are found in these passages.) The best explanation of this textual phenomenon is to suppose that, in Hume's view, to establish that the states produced by a psychological mechanism are beliefs is automatically to establish that they are justified, other things being equal (SJ 33, 65, 73–4). We thus need to locate a property that is *necessary* for the states produced by a psychological mechanism to constitute beliefs, such that to establish that the states are beliefs and thus have this property is also *sufficient* to establish that the beliefs are justified, other things being equal. What could this property be? The only salient candidate is steadiness—itself a species stability, stability in virtue of being infixed by custom or repetition. This argument for taking Hume to connect justification to stability is based on passages in 1.3.5–10, where Hume is advancing a general account of belief that applies to common persons just as well as to the learned. If long-term consensus were the sole epistemic goal, this key evidence for a stability interpretation would go by the boards.

There is a third difficulty. In light of the many passages that foreground the epistemic achievements of common persons, Hume must have on offer some account of this ordinary knowledge. If long-term consensus within the community of inquirers is the sole epistemic goal, Hume would need to construct a general theory of justification in these terms. He would need to assess the epistemic status of common persons in virtue of their standing in an appropriate relationship to that of the learned. On this approach, the vulgar possess justification derivatively—perhaps to the extent that their beliefs approximate those contained in the long-term consensus; or to the extent that their beliefs result from mechanisms that approximate those licensed by the goal of long-term consensus; or to the extent that their beliefs and/or belief-forming mechanisms causally contribute to long-term consensus, or have it for a natural result. In each case, the epistemic standing of the common person is derivative, parasitic on that of the learned. Waive the earlier point that such constructions might yield indeterminate, or determinate but undesirable, assessments. What matters now is that parasitism reverses the direction of explanation that can plausibly be attributed to Hume. In conferring positive epistemic epithets on mundane beliefs, Hume does not so much as pause to suggest that he needs to explain how the short-term epistemic successes of common persons derive from a goal of long-term consensus among the learned.

We cannot interpret Hume as putting forward long-term consensus within the community of inquirers as the sole epistemic goal. It is doubtful that this goal legitimizes the epistemic distinctions Hume wishes to draw. Hume's arguments for the value of stability apply directly to any individual; more generally, central textual evidence for a stability interpretation derives from epistemic remarks about beliefs resulting from mundane causal inference. Further, Hume sees no need to appeal to long-term consensus to explain everyday epistemic accomplishments.

The passages Williams cites as evidence of the character of Hume's epistemology must be weighed against these problems.

Perhaps Hume adopts two epistemic goals, individual stability for common persons and long-term consensus for those engaged in scientific inquiry. Even so, Hume would owe us an account of the connection between these goals. There is an elegant, interpretive alternative: Hume promulgates a single theory of justification, in terms of individual stability, applicable to the vulgar and the learned alike. Williams and I agree that some conception of stability is essential to Hume's theory of justification. I contend that long-term consensus plays no role; the theory is exhaustively characterized in terms of individual stability. We can grant that long-term consensus is a goal distinctive to science and philosophy, the community of inquirers (cf. Williams 284–6, 293), but the epistemic status of those pursuing that goal is subsumed under a broader theory of individual justification.

Further, any consensus within a reformed science is a *byproduct* of the goal of individual stability. Deficiencies in the methods of the common person give rise to more refined methods and thus to the distinctive achievements of the learned. Common persons have considerable epistemic success, but also possess beliefs that arise from mechanisms that are vulnerable to instability. On occasions when instability sets in, the individual will be motivated to relieve the associated uneasiness. Should he locate sufficiently systematic means to this end—by adopting "general rules" (SJ §§ IV.2–3; cf. Schmitt 324n4)—the individual will have refined his belief-forming mechanisms so as to enhance their tendency to produce stable belief. For example, unregulated causal inference leads to conflicts—broadly, inductive inconsistencies—and hence uneasiness. This provides a motivation for subjecting causal inference to higher-order inductive principles that control for the effects of accidental circumstances, helping to avoid rash generalizations and prejudice (cf. Williams 285, 291–2). The refined belief-forming mechanisms tend to produce greater stability. Beliefs that survive the transition to the more refined mechanisms will thus possess greater justification than before. Reliance on belief-forming mechanisms that are most conducive to stability, and hence have the highest epistemic standing, arises (when it does arise) progressively under the pressure of individual discomfort in instability. We are thus led to more sophisticated modes of inquiry. Bacon's theme of "co-operation" (Williams 287)—and institutional arrangements designed to promote collaboration—seems less than prominent. Consensus is achieved on the basis of procedures that are a natural outgrowth, individual by individual, of common epistemic deficiencies.

This overall picture is similar to Hume's strategy in the case of moral judgment. He wants to explain why some judgments of character have normative pride of place and to do so by explaining how moral judgments arise. The trick is to show that the explanation is one of a particular kind, involving psychological processes that result in enhanced stability. Thus, the discomfort of the instability

in the continual "contradictions" and fluctuations motivates the adoption of a steady and general point of view that stabilizes judgments of character. Suppose Hume restricts himself to the relatively sparse resource of *intra*personal variation. In order to steady his own variation in judgment, each individual adopts a point of view that is psychologically salient. The same point of view—that of the agent's narrow circle—is salient to all, so that the judgments of individuals converge. Consensus—convergence—results from a natural motive to relieve individual uneasiness; no one need adopt consensus as an independent objective (parts 2 and 4 of my response to Kelly). We best view Hume's explanation of the adoption of general rules as following a similar strategy—the normatively preferred beliefs arise, under the pressure of instability and uneasiness, from within a less refined, individual perspective.

Hume may have an extensive interest in long-term consensus, but it is not a goal to which he would appeal for the purpose of generating normative epistemic discriminations, not the backbone of his theory of justification. Williams observes: "[F]rom Hume's standpoint, human cognitive powers, while possibly more extensive than those of other animals, are not essentially different from them and are to be studied and explained in the same empirical and decidedly secular spirit" (Williams 268). I would say studied, explained, and *assessed*. Epistemic success extends to non-human animals: "the most ignorant and stupid peasants, nay infants, nay even brute beasts . . . *learn* the qualities of natural objects" (EHU 4.2.23, emphasis added; SBN 39). Thus, both the *Treatise* and the first *Enquiry* devote entire sections to the "reason" of animals—animal cognition as well as belief (part 1 of my reply to Schmitt). A lengthy footnote explains why "men so much surpass animals in reasoning, and one man so much surpasses another" (EHU 9.5n; SBN 107n1), even though all just reasoning derives from custom.[51] Epistemic achievements, in different degrees, are possible for animals, the vulgar, and the learned. I have my own candidate for the most important single theme in Hume's constructive epistemology (cf. Williams 290): the continuity of the epistemic achievements of the most sophisticated inquirers with those of the most humble organisms. This is part of the data stability is invoked to explain.

NOTES

I want to express my deep appreciation to Erin Kelly, Frederick Schmitt, and Michael Williams for their participation in this symposium and the enormous amount of time and thought they have invested in it. I thank Elizabeth Anderson for a briefing on the psychological literature on prejudice, and Donald Ainsele, Stephen Darwall, James Joyce, Peter Millican, Lawrence Sklar, and David Velleman for helpful discussion. I am grateful to the editors of *Hume Studies* for this opportunity for consideration of my work,

and to Kenneth Winkler for his extraordinary care and patience in providing extensive, detailed suggestions. Thanks also to Erica Lucast for help preparing this manuscript.

1 I employ parenthetical references for the symposiasts' papers and for my book: EK—Erin Kelly, "Stability and Partiality in Hume's Moral Philosophy: A Response to Louis Loeb," *Hume Studies* 30 (2004): 329–38; FS—Frederick Schmitt, "Loeb on Stability and Justification in Hume's *Treatise*," 297–327; and MW—Michael Williams, "The Unity of Hume's Philosophical Project," 265–96; SJ—*Stability and Justification in Hume's* Treatise (Oxford: Oxford University Press, 2002).

2 I use the following abbreviations for quotations from, and citations to, Hume's works: Abs.—*An Abstract of* . . . A Treatise of Human Nature, in *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2000); EHU—*An Enquiry concerning Human Understanding*, ed. Thomas L. Beauchamp (Oxford: Clarendon Press, 1999); EPM, *An Enquiry concerning the Principles of Morals*, ed. Thomas L. Beauchamp (Oxford: Clarendon Press, 1998); *Essays*—*Essays, Moral, Political, and Literary*, ed. Eugene F. Miller (Indianapolis: Liberty Classics, 1987); L—*Letters of David Hume*, 2 vols., ed. J. Y. T. Grieg (Oxford: Clarendon Press, 1932); T—*A Treatise of Humean Nature*, eds. Norton and Norton. In the case of the *Abstract*, *Treatise*, and *Enquiries*, page references are also provided to one of the following editions: SBN—*Enquiries concerning the Human Understanding and concerning the Principles of Morals*, ed. L. A. Selby-Bigge, 3rd ed., revised by P. H. Nidditch (Oxford: Oxford University Press, 1975); SBN—*A Treatise of Human Nature*, ed. L. A. Selby-Bigge, second edition, with text revised and variant readings, by P. H. Nidditch (Oxford: Oxford University Press, 1978), which also contains the *Abstract*.

3 Even enemies need not be *stigmatized*. Our sympathetic response to them might nevertheless be less intense, other things being equal, in light of their differences from us, dissimilar concerns and goals. (This is the inverse of heightened sympathetic response, discussed in note 16, resulting from cooperation.)

4 Christine Korsgaard, "Normativity as Reflexivity: Hume's Practical Justification of Morality," Sixteenth Hume Society Conference (Lancaster, England), 1989, introduction and part I, and *Sources of Normativity*, ed. Onora O'Neill (Cambridge: Cambridge University Press, 1996), 62–3. The Hume Conference paper is unpublished.

5 *Sources of Normativity*, 62. Similarly: A "faculty need only give a positive verdict about its own operations. . . . The understanding, generally speaking, does not"; "it does not, at least as long as we engage in . . . reflection" ("Normativity as Reflexivity," introduction and part I). The qualification is also reflected in *Sources of Normativity*, 63n31. Korsgaard thus recognizes Williams's point that the destabilization is temporary; since reflective endorsement is closely related to the "reflective stability" account (SJ 89n46), skepticism ensues nevertheless. See part 3 of my reply to Schmitt and part 1 of my reply to Williams.

6 *Sources of Normativity*, 62–3, and "Normativity as Reflexivity," part I. Korsgaard also finds evidence for a reflective endorsement interpretation in the second *Enquiry* (*Sources of Normativity*, 51–4, and "Normativity as Reflexivity," part I) and in Hume's aesthetics ("Normativity as Reflexivity," part I, n27). For some opposing textual evidence, see SJ 28n47. Korsgaard maintains that in the *Treatise* "Hume appealed to a more specific version of the reflective endorsement account, . . . 'normativity as reflexivity'" (*Sources of Normativity*, 61).

7 Christine M. Korsgaard, "The General Point of View: Love and Moral Approval in Hume's Ethics," *Hume Studies* 25 (1999), 3–41, 14–15.

8 Korsgaard, "The General Point of View," 24–5.

9 There is some ambiguity. Kelly writes: "More precisely, a character trait is picked out as virtuous by the approval that we feel when we sympathetically identify with those persons who would benefit from the agent's disposition to act in the associated ways" (Kelly 331). If "sympathetically identify with" means *sympathize* (in Hume's technical sense of the term) *with* the experiences of the narrow circle, this yields my reading of Kelly. If "sympathetically identify with" means *to adopt the perspective of* a member of the narrow circle, that is, to sympathize *from* their point of view, Kelly's interpretation is equivalent to the one I give below.

10 Compare Terence Penelhum's suggestion in *Hume* (London, Macmillan, 1975), that "objectivity," for Hume, involves "imagined closeness" (142–3). The account, however, does not fully succeed. A more consistent development of a Humean position would identify the steady point of view with that of the agent himself! See my "Hume's Agent-centered Sentimentalism," *Philosophical Topics* 31 (2004): 309–41, § III. (I do not want to rely on this paradoxical result in addressing the issues Kelly raises.) This article also elaborates other themes in this part of my response to Kelly.

11 There are perhaps cases where a person is *de facto* deprived of a narrow circle, but it is plausible that the general rules specifying the usual circumstances associated with a character trait include the typical composition of an agent's narrow circle. On this point, see Charlotte Brown, "From Spectator to Agent: Hume's Theory of Obligation," *Hume Studies* 20 (1994): 19–35, 24; cf. 25.

12 The resulting normative theory is similar to C. D. Broad's "self-referential altruism." See "Self and Others," in *Broad's Critical Essays in Moral Philosophy*, ed. H. D. Lewis (London: George Allen & Unwin, 1971), 277–82. Rico Vitz takes note of this affinity in "Hume and the Limits of Benevolence," *Hume Studies* 28 (2002): 271–95, 284–5, including notes 11 and 12. Vitz cites Broad's "Moore's Ethical Doctrines," in *The Philosophy of G. E. Moore*, ed. Paul Arthur Schilpp (Northwestern University: Evanston and Chicago, 1942).

13 These issues are rarely discussed in connection with the steady point of view. For a notable exception, see Kate Abramson, "Correcting *Our* Sentiments about Hume's Moral Point of View," *Southern Journal of Philosophy* 37 (1999): 333–61, esp. 345, 347–50.

14 Kelly's examples—prisoners held without legal rights, zoning restrictions—involve systems of law and statute. So too with the cases of prisoners in war; Hume contemplates laws of nations and laws of war in particular (T 3.2.11.1; SBN 567, EPM 3.1.11, 4.2; SBN 187–8, 205). We need to ask how Hume's account of justice might speak to these examples (cf. Kelly 333–4). I do not see anything in Hume's normative theory to preclude non-trivial obligations to groups suffering from or at risk of injustice.

15 One reason the agent's narrow circle constitutes a salient sympathetic perspective may be that information about the agent is more readily available within than outside the circle. Korsgaard gestures in this direction ("The General Point of View," 15).

16 There is substantial empirical evidence in the tradition of Gordon Allport, *The Nature of Prejudice* (Cambridge, Massachusetts: Addison-Wesley, 1954), that contact subverts

prejudice. For one recent development, emphasizing contact in the cooperative pursuit of a common goal, see Samuel L. Gaertner and John F. Dovidio, *Reducing Intergroup Bias, The Common Ingroup Identity Model* (Philadelphia: Taylor & Francis, 2000).

17 I am taking some liberty, substituting *family* for *relations* in Hume's formulations. This sidesteps complications in that relations need not be nested (see below) relative to the other categories.

18 In "Loyalties," *Journal of Philosophy* 79 (1982): 173–93, Andrew Oldenquist develops a moral theory that is "neither egoism nor impersonal morality" (176). The theory is a version of Broad's self-referential altruism (n12 above), of weighted universalism. Interestingly, Oldenquist focuses on loyalties and obligations to groups that are nested by set inclusion (178, 180–1). Hume's thought that a psychological nesting coincides with this ordering provides a theoretical basis for Oldenquist's selection of groups. The same point applies to Broad, who notes that "Each person may be regarded as a centre of a number of concentric circles," "innermost" to "outermost" ("Moore's Ethical Doctrines," 55). Broad might be read as suggesting a different rationale: we are involuntarily "born into" (55) family and country. This would not explain why friends constitute one of the circles.

19 Some commentators add such categories as "colleagues" or "co-worker" in their accounts of the narrow circle. See Brown, "From Spectator to Agent," 25, and Korsgaard, *Sources of Normativity*, 55, and "The General Point of View," 3. The inclusion of such categories is within the letter of Hume's characterizations of the narrow circle as the "sphere" in which "any person moves" (T 3.3.3.2; SBN 602, cf. 603). Hume nevertheless takes the steady point of view to abstract from such categories.

20 The specification of the usual circumstances may depend upon the kind of character trait under evaluation and upon the kind of agent who possesses it. See Abramson, "Correcting *Our* Sentiments," 343–4.

21 Stephen Darwall has pressed this matter with me. This may be a special case of issues raised in the preceding note, but it is one that deserves special mention.

22 "[N]othing but *a more vivid and intense conception of an idea*" (SBN 119–20). In this paragraph, Hume twice characterizes belief in terms of "force," language that can be used to refer either to vivacity or steadiness (SJ 66). I offer an explanation of why Hume's official theory of belief diverges from the steadiness conception that other passages demand (SJ 66–7, 224–5).

23 The reference to reliability does not commit me to attributing a reliability theory of justification to Hume (SJ 83n40).

24 For this reason, contrary to Schmitt (Schmitt 325n12), a comparative stability interpretation (SJ 96, 110) perhaps does find a rationale in the value of stability.

25 Schmitt seems not to track my stipulation when he says that "the reflection to which the [reflective stability] account refers generally falls short of sophisticated Humean philosophical reflection" (Schmitt 307). I am at fault here.

26 In her contribution to an American Philosophical Association "Author Meets Critics" session (December 27, 2003), Janet Broughton offers some interesting objections to interpreting Hume's project in this way.

27 My presentation here differs from that in the book, where I say that on the reflective stability account "no belief is justified" (SJ 91; cf. 12, 14, 16, 89, 90), thus requiring that "Hume abandons his pretheoretical epistemic commitments" if he is to retain the account (SJ 91; cf. 15). These claims are too strong in light of versions of the reflective stability account that I fail to consider. (See the next part.) Putting to the side the question of whether intense reflection is instrumental to stability (Schmitt 307; Williams 289–92), these other versions do require Hume at least to lower or discard some of his highly positive evaluations. The variant of the ratification argument in the text takes account of this.

28 Williams can offer an alternative explanation: Hume does not relinquish or modify his positive evaluation even of the beliefs of the intensely reflective person. I take up this matter in part 3 of my reply to Williams.

29 I do write of Hume's "contradictory assessments" (SJ 14, 16) of various epistemic claims in the positive and the negative stages, but I issue "an alert to the preliminary character of the discussion in the present section [§ I.4]" (SJ 18), which is introductory.

30 At a late stage in my discussion, I muster a gesture in the direction of the temporal stability account—"Hume could allow that reflective persons are justified during unreflective periods by tinkering with the time frame for stability" (SJ 100).

31 Schmitt's idea is that Hume shows in 1.4.4 that the belief in matter is false (Schmitt 315). The bearing of this result on double existence, however, is unclear; the second set of objects is a set of perceptions, not matter. See part 2 of my reply to Williams.

32 There are issues about whether stability is to be evaluated in terms of belief tokens or types (Schmitt 325n14). My book points toward the latter option, without explicitly embracing it. I do not pursue these issues here.

33 Granted, at step (c) of the argument at 1.4.4.7, Hume claims that the belief in the dependence of accidents arises from custom, but this is not to employ causal inference to evaluate that belief, much less to evaluate the belief specifically in substance.

34 He develops this interpretation in *Knowledge and Belief* (London: Routledge, 1992), chap. III, with special reference to 1.4.2. I am unsure how Schmitt would apply reliabilism to Hume's account of moral judgment.

35 In calling attention to the most promising "method that can close the interminable debate over philosophical issues," Schmitt moves in Williams's direction (Schmitt 326–7n19). But Schmitt wants to distinguish between the aim of stability, which applies to philosophical inquiry, and the aim of truth germane to "cognition that influences the will, to which the status of justification applies." I am uncomfortable with this proliferation of goals. See part 4 of my reply to Williams.

36 In her APA comments, Broughton shrewdly observes that my Hume does not "explain what aspects of our experience and our minds give rise to the concept of justification," whereas he does explain the origins of other concepts or kinds of discourse. It would be well to explain this.

37 For discussion of the class of externalist interpretations, and some other themes in this part, see my "Inductive Inference," forthcoming in Elizabeth Radcliffe, ed., *The Blackwell Companion to Hume*. I provide an internalist variant of a stability account in

"Integrating Hume's Accounts of Belief and Justification," *Philosophy and Phenomenological Research* 63 (2001), 279–303.

38 For affinities between Hume's and Peirce's theories of the nature and fixation of belief, see SJ § III.6 and 124n34.

39 Jane McIntrye advances a version of this criticism in her "Comments On Loeb's *Stability and Justification in Hume's* Treatise," 31st Annual Hume Society Conference, August 2–7, 2004, Tokyo.

40 Even the interpretation in my book does not take it *all* away; justification remains for persons who are not fully or hyper-reflective (part 2 of my reply to Schmitt).

41 The temporal stability variant thus has affinities with contextualist theories; elsewhere, Williams has suggested an interpretation along contextualist lines. See Schmitt 324n8 and SJ § III.7.

42 I am less sure. There is no mention of cause and effect reasoning in the double vision argument; it is fundamentally an argument by parity of case.

43 The considerations, however, are different in the two cases: whereas there is no non-circular *enumerative* argument for the uniformity of nature, there is no legitimate *theoretical* inference to a second set of objects.

44 Confused, rather than false—conceptual skepticism. See my "Hume's Explanations of Meaningless Beliefs," *The Philosophical Quarterly* 51 (2001): 145–64, esp. § V.

45 The imaginative propensities include the identity propensity and the mind's propensity to spread itself on objects. (Cf. Schmitt 324–5n9.) This latter propensity leads to the beliefs in local conjunction and in necessary connection. See "Hume's Explanations of Meaningless Beliefs," esp. §§ IV–V.

46 John Passmore pursues this theme in *Hume's Intentions* (London: Duckworth, 1952/1968), 14–15 and 151. Hume's hopes are dashed in the Appendix (T App.10; SBN 633).

47 I make this point too cryptically in the book. For a fuller development, see "Hume on Stability, Justification, and Unphilosophical Probability," *Journal of the History of Philosophy* 33 (1995): 101–32, 107–11.

48 See Passmore, *Hume's Intentions*, 150–1.

49 It is impressive that both Williams and Schmitt (Schmitt 301–2, 307) find evidence that intense reflection is instrumental to stability, but interesting that they extract the doctrine from different portions of the section. Schmitt's evidence, from 1.4.7.10 (SBN 269), is within a stage where Hume is describing "sentiments of my spleen and indolence" (T 1.4.7.11; SBN 270). This phase closes out at the end of the eleventh paragraph (cf. SJ 87n45).

50 Williams writes that "Hume is notable for his avoidance of the first person. He prefers to speak of 'we': we the learned" (Williams 285). Yet, consider the chief tandem passage, at 1.3.9.4 (SBN 108), material on which Williams himself relies (Williams 276–7).

51 Schmitt's reliabilist interpretation also does well by these themes. For elaboration, see "Some Context" and "Hume's Epistemic Options" in my "Inductive Inference."