Midwest Studies in Philosophy, XXXI (2007)

The Epistemology of Thought Experiments: First Person versus Third Person Approaches

KIRK LUDWIG

There has been a movement recently to bring to bear on the conduct of philosophical thought experiments (henceforth "thought experiments")[1] the empirical techniques of the social sciences, that is, to treat their conduct as in the nature of an anthropological investigation into the application conditions of the concepts of a group of subjects. This is to take a third person, in contrast to the traditional first person, approach to conceptual analysis. This has taken the form of conducting surveys about scenarios used in thought experiments.[2] It has been called "experimental philosophy" by its practitioners and has been applied across a range of fields: the philosophy of language, the philosophy of action, the philosophy of mind, epistemology, and ethics.[3] The results of these surveys have been used to support conclusions about the application conditions of particular concepts of interest in philosophy. They have also been used to support (and been motivated by) skeptical claims about the traditional approach to conceptual analysis. The

1. I mean to distinguish philosophical thought experiments, which aim at the elucidation of concepts, in a sense to be further explained, from scientific thought experiments, whose aim is to draw out consequences of empirical theories of the world.

2. Many of the themes expressed by proponents of "experimental philosophy" are similar to those expressed by Arne Naess in his early work, especially in *Interpretation and Preciseness* (1953), which presents a wealth of empirical information about ordinary usage in the form of surveys, along with analysis.

3. See, for example, Knobe (2003a,b, 2004, 2005, 2006); Machery et al. (2004); Malle and Knobe (2001); Nadelhoffer (2004a,b, 2005); Nahmias et al. (2004, 2005); Nichols et al. (2003); Nichols (2004a,b); and issues 1 and 2 of volume 6 of *The Journal of Cognition and Culture*, 2006.

purpose of this article is to evaluate the usefulness and limitations of the methodology of surveys as a way of pursuing the traditional goals of conceptual analysis in comparison with the more traditional "armchair" method and the skeptical claims made on the basis of survey results about the traditional method of conceptual analysis as a method of attaining knowledge.[4]

I will argue that there is no good ground for skepticism about the traditional first person method of philosophical analysis to be found in the results of surveys of ordinary people about philosophical thought experiments, that while surveys have some interest, they do not represent a fundamentally new way of doing philosophy, they are liable to considerable methodological difficulties, and that they cannot be substituted for the first person method but must eventually lead back to first person assessments of thought experiments, as these are crucial to acquiring the first order knowledge and understanding which is our object in philosophy, and which is the only standpoint from which to acquire *a priori* knowledge.

In section 1, I give an account of the traditional "first person" method of conceptual analysis and the conduct of thought experiments. This will provide the framework I will use in evaluating the relevant implications and merits of the first versus third person approach to thought experiments. It will not be part of my purpose to defend the framework,[5] for the challenges to it which I will consider do not the take the form of a request for a foundational justification of it but the form of the claim that its traditional application from the first person standpoint as primary is flawed or not fully adequate to the tasks philosophy sets itself. In section 2, I review two examples of the use of the "experimental method" in philosophy to serve as foils for discussion. The first aims to cast doubt on the tenability of the traditional enterprise of conceptual analysis by appeal to systematic differences in responses to thought experiments across cultures, specifically with respect to questions about what fixes the referents of proper names. The second aims to pursue the traditional aims by the survey approach with respect to the question of the conditions under which it is correct to say that someone did something intentionally under a certain description. My aim throughout is to focus on general methodological issues rather than on resolving the particular philosophical issues under discussion. In section 3, I turn to evaluating whether the results of the cross-cultural surveys cast doubt on the traditional enterprise and more generally what advantages and disadvantages the third person method has vis-à-vis the first person method. Section 4 is a brief summary and conclusion.

1. CONCEPTS, CONCEPTUAL ANALYSIS, AND THOUGHT EXPERIMENTS

Concepts correlate at the level of thought with word meaning. Words express concepts, and two words express the same concept if and only if (iff) they are

4. Some question arises about whether indeed experimental philosophy is an enterprise intended by at least all its practitioners to have the sort of connection with conceptual analysis as traditionally conceived suggested here. See note 26 for discussion.

5. Though for responses to the so-called calibration problem for rational intuition, see Foley (1998) and Sosa (1998).

semantically complete and synonymous. Concepts are thought of as playing in the content of thoughts a role analogous to that meaningful words play in declarative sentences. When a word is used in the complement sentence in a propositional attitude report, the thought attributed is said to involve or deploy the concept, and traditionally the subject would be said to have or possess the concept.[6] We may deploy a concept in thought which no word or phrase in our language expresses and on that basis introduce a word to express it, as when we introduce, for example, a new color term. Concept possession is thus distinguished from understanding a word, though in practice conceptual analysis goes hand in hand with the analysis of the application conditions of words which express them, since to give an analysis we must introduce a word that is to express the concept. Concepts can be construed broadly as correlating with the meaning of any incomplete expression derived from a sentence by replacing a meaningful word unit or phrase with a place holder or, more narrowly, as correlating with predicate expressions. In the following I will concentrate on concepts as correlates of meaningful, semantically complete, predicate expressions.

Concepts (in the narrower sense I henceforth assume) *apply to* objects or to n-tuples of objects just as predicates are *true of* objects or n-tuples of objects. Conceptual analysis may be thought of broadly as involving two kinds of activities. The first is providing an account of the application conditions of a concept, in the sense of a set of necessary and sufficient conditions for its application, in terms of other, more basic concepts, in the sense of concepts which can be grasped (had or possessed) independently of the concept being analyzed. A standard form for such an analysis can be thought of as a biconditional of the form [C] where "F(. . .)" is a predicate expressing the concept (I return in a moment to remark on the fact that [C] does not mention the concept being analyzed).

[C] For any x, y, z, . . . F(x, y, z, . . .) iff . . .

An instance of [C] constitutes a reductive analysis of the concept expressed by "F(. . .)" iff (1) " . . . " on the right-hand side of [C] is replaced by a formula employing only predicates or expressions understanding of which does not presuppose understanding of "F(. . .)", (2) [C] is analytically necessary, and (3) [C] is concept-giving in the sense that someone who understands the sentence form and the right-hand side and knows that [C] is analytic is thereby able to grasp the concept expressed by "F(. . .)".[7] The second sort of activity is the articulation of the connections between the application conditions of the concept being analyzed and those of other concepts than reductive analysis. This can involve a set of analytically necessary and sufficient conditions for the application of the concept,

6. Those who accept Tyler Burge's (1979) arguments for social externalism will want to relax this requirement, as they allow propositional attitude attributions to be correct in circumstances in which the subject does not possess all the concepts expressed in the complement sentence.

7. For some purposes it may be desirable to impose some additional constraints such as that the analysans be the conceptually simplest concept giving condition that otherwise is satisfactory, or at least as simple as any other, or that it express how our underlying competence in its deployment works, if full sense can be given to that.

though not in terms of concepts which can be grasped independently, or a set of necessary conditions, or a set of sufficient conditions, or an account of what counts analytically as evidence for the application of the concept or what its application counts analytically as evidence for. I will use the term "conceptual analysis" to cover both sorts of activities and the term "reductive analysis" to cover the first sort. Not all concepts of course admit of reductive analyses on pain of infinite regress, and perhaps relatively few do. In this case, analysis comes to the articulation of conceptual connections.

[C] is a material mode statement. It is not a metalinguistic statement, and it does not mention the concept being analyzed or the concepts in terms of which the analysis is given. When we give an analysis of a concept in the material mode, as in [C], it is about the nature of the thing to which the concept applies. This is connected with the object of conceptual analysis, which is to learn something about what it is we use our concepts to think about. This point will emerge as central in our understanding of the importance of the first person approach to thought experiments. It follows that to give an analysis one must understand the sentence that expresses it, and therefore grasp, which is to say, possess, the concepts involved in it.[8] One cannot, therefore, give an analysis of a concept which one doesn't possess.

To possess a concept is to be in a certain sort of epistemic state, one that puts one in a position to articulate its application conditions. The traditional paradox of analysis is how there can be an enterprise of conceptual analysis in the first place if to give an analysis of a concept (word) one must already grasp (understand) it. For if one grasps it already, it seems that one must be in possession of the analysis, and if one does not grasp it, one is not in a position to say anything about its possession conditions. To put it in linguistic terms: How can one be expected to give a definition of a word unless one understands it, and if one understands it, must it not be that one already knows its definition? The form of the solution is to distinguish two different kinds of knowledge, explicit knowledge of the analysis and a kind of implicit knowledge. The implicit knowledge we have of the application conditions of concepts, however, is not propositional knowledge. It is rather expressed in a skill we have in deploying the concept appropriately.[9]

8. I distinguish competence in the use of a sentence from understanding it. Many people pass for competent in the use of sentences which contain expressions they don't fully understand, like "hollyhock," "kosher," "loblolly," and "psoriasis." They are counted as competent because they have enough of a grasp on the normative use to talk sensibly using them in many contexts, even though they can't pick out the hollyhocks in the garden, or tell you precisely what it is for something to be kosher, or tell a loblolly from a southern longleaf, or psoriasis from eczema.

9. It is a misunderstanding to take concept possession to be possession of a propositional theory involving it (Cummins 1998; Ramsey 1998). One can possess a concept without having any word that expresses it and without having any description of it. So the theory would have to involve beliefs which deployed the concept, and this would mean that it would presuppose that one possessed the concept involved in the theory. One's possession of the concept would not then be explained by one's having the theory. Perhaps having some beliefs or some range of beliefs involving a concept is a requirement on possessing it, but what it is to possess it requires in addition the ability to appropriately deploy it in contexts which go beyond any such range of beliefs.

The issue comes into sharper focus if we think about knowledge of word meaning, for when we learn a language we are not typically taught explicit definitions or meaning entailments. Rather, we learn how to use words in relation to things in our environments and in inferences, and we learn what mistakes we make are treated as mistakes of language as opposed to mistakes of fact. We acquire a competence in the use of words in our first languages, and this expresses itself in our being able to make and to be justified in making correct judgments using them in response to information about how things are or might be. This competence then can be exploited in articulating a set of application conditions for predicates by our asking ourselves questions, in the formal mode, about whether the predicate applies, fails to apply, or whether it is left open, relative to various circumstances, actual and imagined, where we rely solely on linguistic competence (in isolation from knowledge of matters of fact) in answering relative to the situation in question. This competence is at the same time (for the relevant class of expressions) a competence in deploying the concepts which the words express.

The knowledge of the application conditions of words we arrive at by exploiting our competence in the use of them is knowledge of the meanings of expressions in a public language and so empirical knowledge of the language. As the words whose application conditions we thereby set out express concepts, we can arrive also at an analysis of the concepts which those words express in the public language. So expressed, this also empirical knowledge, since it is of the form, "the concept expressed by W in L applies to . . . iff _____," where "L" picks out the public language, knowledge of whose semantic properties is empirical.[10]

Using one's own competence in this way to arrive at knowledge of public meanings depends upon knowing that one's competence aligns with the norms of the public language. There are as good reasons to think this is so as there are to think that one inculcated the habits of other members of one's linguistic community (or communities) in learning one's languages, but where questions arise, further checking against the practice of relevant others in the community is called

10. A word is in order about natural kind terms. In my view, competence in the use of natural kind terms does not amount to grasping the concepts they express. How can this be? As "concept" was used traditionally, an object falls under the concept of F-ness iff it has the property of being F, and grasp of the concept of F-ness suffices to know what the property of being F is, that is, what condition an object has to meet to be F without qualification. Since competence in the use of a natural kind term "F" does not put one in a position to know what condition an object has to meet to be F without qualification, it does not suffice for grasping the concept expressed by the predicate "is F." Just as empirical investigation of the world is required to find out what property a natural kind attributes (an investigation guided by the linguistic practices of the relevant community), so it is required to discover what concept is expressed. If the concept expressed is taken to correlate with the meaning of the predicate in the language, as I have assumed, the same goes for the meaning of the predicate. Natural kind concepts, those expressed by natural kind terms, then, are peculiarly unsuited for elucidation by the method of thought experiments couched in terms of the natural kind terms that express them, except insofar as we have empirically discovered what concepts they express, and can rely on that further knowledge beyond ordinary linguistic competence in responding to them. What of intuitions involving Twin Earth and similar scenarios that have motivated the view that it is an empirical matter what properties natural kind terms pick out? These are generated by competence in the use of the words relating to, in effect, how their meanings are to be fixed by how things contingently turn out.

for. In these cases, one can fall back on claims about one's own idiolect, thought of as constituted by one's own dispositions to use the expressions in one's vocabulary (with allowances for words one's understanding of which one recognizes to be at best partial).

We move from empirical to nonempirical knowledge when we move from formal mode claims about the application conditions of words in one's language or claims about the application conditions of concepts expressed by words in one's language to material mode claims which deploy the concepts one is concerned with. An investigation into the application conditions of one's own concepts may be conducted entirely in the material mode. What one then expresses in a standard form analysis of a concept, using one's own idiolect, is not about the public language or about concepts but about what the words are about. If one arrives at it via judgments which express one's competence in the deployment of the concepts involved in it, so that one's acceptance of the analysis is itself an expression of one's competence in deploying the concepts involved in it, then one's knowledge of what it expresses is nonempirical, and in one traditional use of the expression, *a priori*.[11]

A conceptual analysis expresses a conceptual truth. This is not to say, as noted, that it expresses a truth about concepts, but rather that it expresses a truth knowledge of which is attainable on the basis of competence in deploying the concepts involved in it. More generally, a conceptual truth is any truth knowledge of which is attainable on the basis of competence in deploying the concepts involved in it.[12] We may know a conceptual truth and know it *a priori* without knowing that it is a conceptual truth or knowing that it is *a priori* or knowing that it is necessary, if it is. In this respect, judgments that constitute knowledge of conceptual truths are like judgments that express what one remembers.

The role of thought experiments in philosophy is to draw out the implicit knowledge we have of the application conditions of our concepts as it is embodied in our dispositions to deploy words expressing them. Even ordinary judgments about contingent matters of fact, such as that there is a lamp on the desk or that it is raining, express our competence in deploying concepts, for they call on our competence in determining whether the application conditions for the concepts involved in the judgments are met or are likely to be met relative to our perceptual representations of the world around us and our beliefs. Reflection on ordinary

11. One skeptical strain about whether conceptual analysis yields *a priori* knowledge rests on failing to distinguish claims about particular individuals' concepts from claims made using concepts whose warrant is derived from their expressing the subject's competence in deploying them. The first sort of claim is *a posteriori*, but the second sort, for the subject, is not. Conflating them amounts to a kind of use-mention confusion. See Horgan and Graham (1998).

12. Are conceptual truths necessary truths? There are some apparent counterexamples, such as the truths expressed by "All actual philosophers are philosophers" and "dthat(the president of the United States) = the president of the United States." We seem to be able to know these *a priori*, but prepending "it is necessary that" to the sentences expressing these propositions yields falsehoods. In these cases, however, knowledge of the linguistic vehicle used to express the proposition turns out to play a role in what we know, which casts doubt on the claim that what is known *a priori* is the proposition expressed in each case. But in the present context, I will leave this question to one side. Though some are, I make no claim here that all conceptual truths are necessary truths.

judgments we make and the circumstances and why we make them then can help us to determine what the application conditions of our concepts are by the application of Mill's methods: Features which are not present in all the cases (as we understand them) in which we judge that something is F are not necessary for a thing to be F, and features present in all such cases are candidates for necessary and sufficient conditions. Actual judgments we make involving a concept (or word) are finite in number and so the evidence which we accumulate in this way is limited, and our judgments are often also mediated by empirical belief. Since competence in the deployment of concepts puts us in a position to make such judgments about so far unencountered circumstance types, we can extend our evidence by asking whether judgments deploying them would be true in various imagined circumstances, and we can ask with the intention that the answers be based solely on the concepts deployed in the judgment in response to the scenario. When we do so, we are engaged in a philosophical thought experiment.

This is not a matter of asking what we would *say* about the scenario but about what is so with respect to it. And it is not a matter of asking what we would say or think *were we in* the circumstances described, for the circumstances need not involve any thinkers, and even if it involves us, the question is not what we would think were we in such circumstances but what is so with respect to it in regard to the relevant concepts we actually employ.

A thought experiment involves two distinct roles, that of the experimenter, the person conducting the thought experiment, and that of the subject, the person on whom the thought experiment is run. The experimenter presents a scenario to the subject, and then asks one or more (typically material mode) questions about it, involving the concept or concepts under investigation, the responses to which are to be given solely on the basis of the subject's competence in deploying the concepts involved in the questions. Questions need not be restricted to yes–no questions. I will, however, focus on yes–no questions, which ask whether the target concepts apply (in material mode questions) in the scenario. In the case of yes–no questions, four sorts of responses may be given: (1) yes, (2) no, (3) the scenario presents insufficient information to answer the question,[13] and (4) a presupposition of the thought experiment relevant to one of the answers (1) to (3) being correct is false.

Thought experiments are usually conducted under certain idealizations about language, namely, that the terms used to describe scenarios and to ask questions and which are to express the target concepts are all semantically complete and consistent. However, many natural language predicates are vague or in other respects semantically incomplete or, alternatively, inconsistent, in the sense that the rules extractable from the practice for their use enable one to derive contradictions (as in the case of the semantic paradoxes). On a Fregean view of

13. I am indebted here and in the next two paragraphs to Kathy Kanuck Lammen's dissertation, "A Defense of Conceptual Analysis and Thought Experiments" (Ph.D. dissertation, University of Florida 2005).

concepts, these predicates do not express concepts at all.[14] I will follow the Fregean usage because it allows us to frame issues more clearly. When we conduct thought experiments using semantically incomplete expressions, we do not get information about the concepts expressed by them, as there are none, but rather information about our meaning giving practices with respect to them. The fact that the expressions are semantically incomplete, however, means that relative to some scenarios our competence in their deployment may not put us in a position to answer a question about it even if all the information that could be relevant to its application has been provided.[15] The proper response in this case is to say not that the scenario provides insufficient information, but to step back from our idealization about the term in question and note that our semantic practices with respect to it provide insufficient guidance or raise difficulties that prevent any straightforward answer.

When a subject of a thought experiment responds to a question about a scenario, it is important that he or she understand that the response is to be based on his or her competence in deploying the concepts expressed in the question. Conducting and being the subject of a thought experiment is a reflective exercise. It requires that both the experimenter and the subject understand what its point is. As it is a reflective exercise, it also presupposes that the subject of the thought experiment is able to distinguish between judgments solely based on competence (or recognition of the limits of competence) in deploying concepts in response to the described scenario.

It is in this connection that we speak of intuitions or intuitive judgments in response to scenarios.[16] For terminological clarity, I will use "intuition" to mean an occurrent judgment formed solely on the basis of competence in the concepts involved in response to a question about a scenario, or simply an occurrent judgment formed solely on the basis of competence in the concepts involved in it (in response, we might say, to the null scenario).[17] Intuitions in this sense are what we

14. Frege says, "A definition of a concept must be complete; it must unambiguously determine, as regards any object, whether it falls under the concept. . . . If we represent concepts in extension by areas on a plane, this is admittedly a picture that may be used only with caution, but here it can do us good service. To a concept without sharp boundary there would correspond an area that had not a sharp boundary-line all round, but in places just vaguely faded away into the background. This would not really be an area at all; and likewise a concept that is not sharply defined is wrongly termed a concept. Such quasi-conceptual constructions cannot be recognized as concepts by logic; it is impossible to lay down precise laws for them" (Beaney 1997, 259).

15. The fact that most of the terms of ordinary language and science are semantically vague presents deeper problems for our ordinary procedures than are typically recognized, but I will leave these worries aside here. See Ludwig and Ray (2002) for an intimation of the problem.

16. I am indebted here and in the next paragraph to Ivana Simić's work in her dissertation project, "Intuitive A Priori Knowledge: Reliability and Rationality," University of Florida.

17. It is clear that some criticisms of the traditional role of intuitions in philosophy are based on a misunderstanding of how the term is traditionally used in philosophy. Two excellent examples of this are Shafir (1998) and Wisniewski (1998). Both clearly have in mind by "intuition" what people say or judge about something on being asked or reflecting on it, with no attention to whether the matter is clearly empirical or not. The examples they use to show that ordinary "intuitions" are unreliable have to do with what people are inclined to believe or say about what are clearly empirical propositions, such as what people will do in various choice situations and how they reason. These are not intuitions in the sense isolated in the text. Gopnik and Schwitzgebel

seek to elicit in response to questions about scenarios in thought experiments.[18] I will use "response" as a neutral term for a judgment made in response to a question about a scenario. This leaves open that not all responses are intuitions or express just competence with the concepts involved in response to the content of the scenario.

Intuitions often seem to have associated with them a distinctive phenomenology, though I make no commitment about this being necessary. However, if there is a distinctive phenomenology associated with intuition, it is important to note that neither it nor the recognition of it is what warrants or justifies the

(1998, 77) say explicitly: "We will call any judgment an intuitive judgment, or more briefly an intuition, just in case that judgment is not made on the basis of some kind of explicit reasoning process that a person can consciously observe. Intuitions are judgments that grow, rather, out of an underground process, of whatever kind, that cannot be directly observed." They include under this heading judgments about such things as "Bob's likely reaction to an insult." Clearly, this is not what philosophers have had in mind, though they suppose that it is: "Much philosophical work consists of articulating and elaborating these intuitive judgments" (Gopnik and Schwitzgebel 1998, 77) Foley also treats "intuition" as a term covering all "immediate unreflective judgments" (Foley 1998, 245). Failing to draw a distinction between unreflective judgments based on empirical beliefs and judgments based solely on competence in the deployment of concepts not uncommonly leads to a failure to appreciate the special epistemic status of the latter, the special role that first person investigation of them plays in the acquisition of *a priori* knowledge, and the stability of the judgments which are reached on this basis. Gopnik and Schwitzgebel appear to fall into the trap they have set themselves by chastising philosophers who are making claims about the *boundaries of concepts* for trespassing on the empirical preserves of science. But the boundaries of our concepts fix the subject matter about which empirical inquiry is conducted. Without a clear view of the distinction, empirical inquiry is itself thrown into the confusion of supposing new discoveries about old subjects when all that is going on is that words are being used in new ways. We can compare the difficulties here to those that beset the classical British Empiricists due to their failure to distinguish concepts, sensations, and propositional attitudes, using the one term "idea" to cover all of these.

18. My use of "intuition" here differs from that of Bealer (1998, 2002) and Sosa (1998, 2006). Bealer distinguishes intuitions of the sort relevant to philosophical inquiry, which he calls "rational intuitions," by their contents or at least by how their contents are presented to us, namely, "as necessary." This will let in more than I am allowing, but more importantly it may fail to include some judgments I wish to treat as intuitions, for it is not clear to me that all judgments based solely on competence in the use of concepts must have their contents so presented or have modal notions as part of their content. I also wish to take no stand on whether intuitions are a sui generis sort of attitude, as Bealer suggests. So far as the basic picture I have sketched goes, intuitions can be ordinary beliefs. What matters is not the psychological mode but how it is generated. Sosa likewise singles out the intuitions of interest in part by appeal to a feature of their content: Intuitions in the relevant sense have "abstract propositions" as their contents, relying on an intuitive understanding of what this comes to. Both of these approaches have a clear connection with the idea that intuitions are expressions of competence in the deployment of the concepts involved in them. Sosa is interested primarily in characterizing a relation of an individual to a proposition, namely, its being intuitive for s at t that p. He characterizes it in part in terms of a counterfactual conditional, "if s were to understand that p fully at t, then s would believe that p at t," in part in terms of the proposition that p being abstract, and in part in terms of s satisfying the antecedent of the conditional at the time. An intuition would then presumably be the belief that s comes to have. This would be too restrictive for my purposes, however, since I am concerned with prompted judgments, though in these cases it seems clear that there is no reason to deny antecedent full grasp of their propositional contents.

proposition judged.[19] It is warranted if it is because of its etiology, because it is an expression of the competence one has in deploying the concepts, of a skill in making correct judgments on the basis of grasp of the concepts involved. In this respect, I use "intuition" in a way parallel to the way "memory" is used. To say someone remembers something (focusing on propositional memory) is in part to say something about the etiology of a current state of his which suffices for him to know what he remembers. The phenomenology of memory, to the extent to which there is one, is not what warrants what is remembered. Remembering is an expression of antecedent knowledge. In a similar way, intuition is an expression of antecedent knowledge, though not propositional knowledge. The importance of the phenomenology of intuition emerges in reflection on what we know and how we know it rather than fundamentally in justifying the judgments we make or the beliefs we thereby arrive at. As the phenomenology associated with occurrent propositional memory serves as a sign of the source of a propositional judgment, the phenomenology associated with intuition may serve as a sign that the source of the judgment is competence in the use of the concepts involved. It can seem to one that a judgment is an intuition when it is not, just as it can seem to one that one remembers something when one does not.[20]

Needless to say, much of the difficulty of conducting thought experiments well involves, on the part of the experimenter, choosing useful and adequately described scenarios and good (and well-worded) questions about it, and then, on

19. Intuitions are often treated as bearing to the propositions they are about the relation of sensory experience to perceptual beliefs, that is, a separate source of justificatory evidence. See, for example, Goldman and Pust (1998) and Kornblith (1998). But this is to overlook the difference between what justifies or warrants the judgment and what justifies or warrants a second-order claim about concepts or words.

20. A world about the role of "reflective equilibrium" in conceptual analysis (of both the narrow and broad varieties): A criticism sometimes made of the reflective equilibrium method of arriving at a correct conceptual account is that it provides nothing but an internal justification for the conclusions that it reaches (Stich 1998). It is in this respect thought to be like coherence theories of justification generally, and the difficulty is highlighted by imagining conflicting but equally internally coherent sets of beliefs. The current framework, though, does not advocate a method of reflective equilibrium in this sense. First, intuitions being classified on the basis of their etiology, their reliability is guaranteed. The issue that reflective equilibrium addresses, in these terms, is how to determine what are the real as opposed to seeming intuitions, and it recommends those that fit with most of the others. This is a valid procedure on the assumption that most of the seeming intuitions are in fact intuitions, for then those in conflict with most will be just seeming intuitions. The case is similar with respect to the reflective equilibrium method applied to knowledge of the past, which rests on, as we might put it, memory impressions, understood neutrally between what we remember and what we seem to remember. If most of them are instances of remembering, then the procedure will help us to weed out the mere memory impressions. Similarly, if one can establish, as Davidson hoped to by a transcendental argument, that most of our empirical beliefs are true, coherence with most of our empirical beliefs becomes a well-grounded method for testing belief. But why accept the relevant assumption? In the case of intuitions, it is because to possess concepts and to use them we must in fact be competent in their deployment. Slip-ups will be explained by performance errors or mistakes about what is being asked or what assumptions prompt a response, or by correctable errors in reasoning. But there could be no such thing as massive systematic error, for this would undercut the assumption that we possess the concepts in question.

the part of the subject, distinguishing between judgments based solely on competence in deploying concepts in relation to the scenario as described and, variously, judgments based on well-entrenched empirical beliefs, judgments based on what would be standardly implicated by the sentence that expresses the judgment, judgments based on reading more or less into the scenario than is intended by the experimenter, judgments based on information carried by the linguistic vehicle for it as opposed by what it means, and judgments made in circumstances in which none should be made because the idealization that the terms involved are semantically complete breaks down dramatically, as in a borderline case of "bald" or "heap" or "person," or, generally speaking, judgments whose etiology is not solely the competence of the subject in use of the concepts in response to the scenario as described. Clearly, training, both on the part of the experimenter and on the part of the subject, is important for well-conducted thought experiments. It is important that the experimenter know what the methodology of thought experiments is and what sorts of difficulties arise in conducting them and getting useful results. It is also important that the experimenter have some knowledge of prior work on the conceptual field in which he is doing work and where the problem areas are in investigating the structure of the concepts he is concerned with. The subject of the experiment should likewise be apprised of the methodology of thought experiments so that he knows what task is being set him and what his responses are supposed to be based on. Otherwise, he may well give a response which does not answer to the purposes of the thought experiment. He may think he is being asked to say what people would say, or he may not recognize that he is not to imagine himself in the situation, or to set aside contingent beliefs about how the world is, or to look for biases which may infect his conception of the scenario. Or he may just say the first thing that he thinks without attention to trying to answer on the basis of what the concepts involved dictate. As for other intellectual disciplines, it is possible to develop one's skills in applying the methods which have been developed for the pursuit of philosophical knowledge and, in particular, to develop one's skills in responding to questions about described scenarios on the basis of one's competence in concepts involved, and this is one of the skills which students acquire in studying philosophy.

There is nothing in this description of the method of thought experiments to suggest that the correct response to a question about a scenario will be obvious. Sometimes this will be so. Gettier's counterexamples to the traditional analysis of knowledge as justified true belief are examples. But in many cases it may not be at all obvious what the right response is, even if we think enough information has been given to determine the correct answer, for it may involve complex inferences from basic principles governing the concepts, and it may involve considerable efforts to screen out distracting factors, as is often the case in thought experiments involving normative concepts drawn from ethics.

Finally, the distinct roles of experimenter and subject in the conduct of thought experiments can be occupied by one individual or by distinct individuals. The first sort of case represents the first person methodology in the conduct of thought experiments. The second sort of case represents the third person methodology in the conduct of thought experiments. The first is the traditional "armchair"

approach to conceptual analysis, though the derogatory phrase "armchair" is of course a caricature of traditional practice, which, like mathematics, has a strong first person methodology as well, has always been a community practice (cf. Bealer's understanding of the enterprise of philosophy as a civilization-wide practice (1998, 203)). The most fundamental question about experimental philosophy, which represents a robustly third person, even anthropological, approach to thought experiments and conceptual analysis, is whether it offers any fundamental advantages over the first person approach and whether it involves pitfalls, difficulties, and limitations which are not present in the first person approach.

2. EXPERIMENTAL PHILOSOPHY

In this section, I review two applications of the third person approach to thought experiments. The first employs thought experiments designed to elicit intuitions relevant to how proper names function (Machery et al. 2004). It involves surveys of East Asian undergraduates and Western undergraduates and reveals some statistically relevant differences in their responses, and on this basis concludes that "philosophers must radically revise their methodology" (Machery et al., B9). The second is an investigation by the use of surveys of the question whether an agent intentionally A's only if he intended to A or whether in some cases whether someone intentionally A's depends on whether A-ing is something that he foresees as a consequence of something he intends to do and believes to be in some respect bad.

(a) Cross-cultural surveys and skepticism about the "armchair" method. I will briefly describe the survey tools and the results, then describe the interpretation offered of it by the authors. Two basic scenarios were used. The first is the Gödel scenario and the second is the Jonah scenario, given, in Kripke's words, in [a] and [b].

> [a] Suppose that Gödel was not in fact the author of [Gödel's] theorem. A man called "Schmidt" . . . actually did the work in question. His friend Gödel somehow got hold of the manuscript and it was thereafter attributed to Gödel. On the view in question, then, when our ordinary man uses the name "Gödel," he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description "the man who discovered the incompleteness of arithmetic." . . . But it seems we are not. We simply are not. (Kripke 1980, 83–84)
>
> [b] Suppose that someone says that no prophet ever was swallowed by a big fish or a whale. Does it follow, on that basis, that Jonah did not exist? There still seems to be the question whether the Biblical account is a legendary account of no person or a legendary account built on a real person. In the latter case, it's only natural to say that, though Jonah did exist, no one did the things commonly related to him. (Kripke 1980, 67)

These examples are supposed to tell against the view that a speaker who uses a proper name uses it to refer to the unique individual who satisfies the descriptions

he associates with the name or a weighted most of the descriptions he associated with it. It is one of the background assumptions of the first case that the only substantive description the ordinary man in question associates with "Gödel" is "the discoverer of the incompleteness of arithmetic." In the second case, it is part of the background story that the substantive descriptions that most people who use "Jonah" associate with the name are derived from what is said about him in the Book of Jonah in the Bible.

The survey probes were modeled after these two stories. The surveys were conducted with two groups: 40 undergraduates at Rutgers University and 42 undergraduates at the University of Hong Kong. Nine students were excluded from the Rutgers sample because they were from East Asian backgrounds. One was excluded from the Hong Kong sample. Four probes were used, two modeled on Kripke's Gödel case and two on Kripke's Jonah case. One of each used names that would be more familiar to members of the Hong Kong group. An example of one probe modeled after the Gödel case, which we will examine more closely later, is the following:

> Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Now suppose that Gödel was not the author of this theorem. A man called "Schmidt" whose body was found under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Thus he has been known as the man who proved the incompleteness of arithmetic. Most people who have heard the name "Gödel" are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel. When John uses the name "Gödel," is he talking about:
>
> (A) the person who really discovered the incompleteness of arithmetic?
> or
> (B) the person who got hold of the manuscript and claimed credit for the work?

If a subject answers (A), he is taken to have intuitions consonant with the descriptivist theory of how the referent of a speaker's use of a proper name is determined, and if he answers (B), he is taken to have intuitions consonant with the causal theory of how the referent of a speaker's use of a proper name is determined. In scoring the reports, (B) answers were assigned 1 and (A) answers 0. Each participant responded to a pair of examples and the scores for each of the pair of probes for each participant were summed. The mean scores and standard deviations are given in Table 1.

Table 1. Survey Results

Case	Score (standard deviation)
Gödel cases	
Western participants	1.13 (0.88)
Chinese participants	0.63 (0.84)
Jonah cases	
Western participants	1.23 (0.96)
Chinese participants	1.32 (0.76)

The results indicate a significant difference between the Western participants and the East Asian participants in responses to the Gödel cases, with Western participants significantly more often giving (B) answers. The differences in responses to the Jonah cases, however, were not statistically significant.

The authors describe these results as eliciting "culturally variable intuitions," and say that "Chinese participants tended to have descriptivist intuitions, while Westerners tended to have Kripkean ones" (Machery et al. 2004, B7). The failure to elicit differences on the Jonah style cases they suggest may be due to the "lengthy presentation," which was perhaps "too long and complex to generate interpretable data," or that for "pragmatic reasons" participants, Western and Chinese, are unwilling to say the speaker is speaking of no one, as the descriptivist account would imply (Machery et al. 2004, B7).

In a section titled "The End of Innocence," the authors draw some startling conclusions:

> Although there are many empirical questions left open by the experiment reported here, we think that the experiment already points to significant philosophical conclusions. . . . the evidence suggest that it is wrong for philosophers to assume a priori the universal validity of their own semantic intuitions. . . . The high standard deviations in our experiment indicates that there is a great deal of variation in the semantic intuitions within both the Chinese and Western groups. . . . A more extreme but very live possibility is that the variability exists even at the individual level, so that a given individual might have causal-historical intuitions on some occasions and descriptivist intuitions on other occasions. If so, then the assumption of universality is just spectacularly misguided. (Machery et al. 2004, B8)

In response to the suggestion that

> . . . philosophers might maintain that . . . only reflective intuitions, i.e., intuitions that are informed by a cautious examination of the philosophical significance of the probes, are to be taken into considerations. (Machery et al. 2004, B8–9)

they say,

> We find it *wildly* implausible that the semantic intuitions of the narrow cross-section of humanity who are Western academic philosophers are a more reliable indicator of the correct theory of reference . . . than the different semantic intuitions of other cultural or linguistic groups. Indeed, given the intense training and selection that undergraduate and graduate students in philosophy have to go through, there is good reason to suspect that the alleged reflective intuitions may be reinforced intuitions. In the absence of a principled argument why philosophers' intuitions are superior, this project smacks of narcissism in the extreme. (Machery et al. 2004, B9)

They suggest then that perhaps philosophers ought to be construed as engaged in an empirical enterprise akin to the Chomskian tradition in linguistics, in which case

> . . . our data are especially surprising, for there is little hint in philosophical discussions that names might work in different ways in different dialects of the same language or in different cultural groups who speak the same language. So on this interpretation, our data indicates that philosophers must radically revise their methodology. Since the intuitions philosophers pronounce from their armchairs are likely to be a product of their own culture and their academic training, in order to determine the implicit theories that underlie the use of names across cultures, philosophers need to get out of their armchairs. (Machery et al. 2004, B9).

We will return to these striking claims shortly.

(b) In a series of papers reporting on survey results, Joshua Knobe has argued against what has become known as the Simple View of acting intentionally,[21] namely, that if one A's intentionally, then one intended to A, and one's A-ing came about as a result of one's so intending and roughly in accordance with one's plan.[22] I want to describe a bit of the developing discussion as a focus for raising questions about methodological issues and the relation of the survey method to the first person method of analysis.

It is a natural assumption that if one A's intentionally, then one intended to A. If one intends to A, and carries it out in the way one intended, then one A-ed intentionally. The simplest assumption, then, is that to say that one A-ed intentionally is to say that one A-ed with the intention of doing so (Anscombe [1957, 1] was an early influential proponent). This has been challenged in the philosophical literature by a number of authors (Bratman 1987, 1984; Harman 1976), and defended by others (Ludwig 1992; Adams 1986; McCann 1991). The challenges are driven by reflection on scenarios which are designed to elicit the intuition that

21. Well, maybe not. See footnote 26.

22. This is meant to be only a rough characterization of the conditions under which doing A as a result of intending to do it licenses saying one A-ed intentionally—see Mele and Moser (1994) for a more thorough discussion of the complications in getting it just right.

an agent did something intentionally though he did not intend to do it. One sort of case (though this is not the only kind of case used for these purpose—see Bratman's 1984 ambidextrous video game player case in particular) focuses on responses to the question whether an agent A-ed intentionally when he did not intend to do so but foresaw that what he intended to do would be sufficient for his A-ing and believed the result to be in some respect bad. For example, Gilbert Harman describes a case in which a sniper knows that firing his rifle will alert the enemy to his position and does not want to do this, but considerations in favor of shooting outweigh the disadvantage this presents. Harman says that, in these circumstances, the agent alerts the enemy intentionally, though he did not intend to do so.

The contribution of experimental philosophy to this debate consists in conducting surveys of nonphilosophers about scenarios designed to test how people will respond to scenarios similar to the one Harman describes, but especially where a foreseen outcome is morally problematic (Harman's suggestion was that an agent's having a reason against doing something foreseen but unintended was sufficient for it to be done intentionally). Knobe (2006) summarizes much of the experimental work on the subject. The explicit object of the work is to uncover the "competencies underlying our folk-psychological concept of intentional action" (Knobe 2006, 205). A starting point are the responses that are given in surveys to the following scenarios (distinguished by whether one chooses "harm" or "help" as the relevant verb):

> The vice-president of a company went to the chairman of the board and said, "We are thinking of staring a new program. It will help us increase profits, but it will also harm/help the environment."
> The chairman of the board answered, "I don't care at all about harming/helping the environment. I just want to make as much profit as I can. Let's start the new program."

In a controlled experiment (Knobe 2003a), eighty-two percent of the subjects responding to the harm scenario said that the chairman harmed the environment intentionally, but only twenty-three percent responding to the help scenario said that the chairman helped the environment intentionally. The effect is robust (Knobe 2003b, 2004; Knobe and Burra 2006b; Nadelhoffer 2004a).

Various suggestions have been made about what explains these results compatibly with the Simple View, and it will be useful to review some of them and then some of the experimental work undertaken in response (I follow Knobe 2006 here).

(1) One suggestion (Adams and Steadman 2004b) is that the effect is due to people not wanting to pragmatically imply that the agent in question is not blameworthy, which they would do if they asserted without qualification that the agent did not A intentionally, an implicature which arises because people will frequently excuse themselves or others by appeal to their not having done something intentionally. When given the choice between saying the chairman harmed the environment unintentionally or intentionally, they avoid the first choice because of its

standard conversational implicature. Against this, Knobe has argued that (a) people will not say someone did A in order to B unless they are willing to say that the person A-ed intentionally, but on conducting surveys we find that (b) people tend to say that the chairman harmed the environment in order to achieve a certain goal (Knobe 2004). Adams and Steadman (2004a) have argued that the same pragmatic factors are at work here since if we admit that someone did not do so and so in order to do another thing, we will suggest that he did not do it intentionally, and this will implicate that he is not to blame. So the additional data does not resolve the matter in favor of the rejection of the Simple View. Knobe responds that such "a complex chain of reasoning could hardly take place in the few seconds it normally takes people to answer these questions" (2004, 217).

(2) Another suggestion (Malle and Nelson 2003) is that people first recognize the agent is to blame for A-ing, and then feel pulled to call the agent's A-ing intentional in order to justify their judgment. Knobe suggests an alternative explanation in which subjects recognize A-ing as bad and foreseen and infer from that both that it is an intentional A-ing and that it is blameworthy. To test which model is correct, he conducts an experiment with a scenario which is intended to be one in which (i) an agent A's knowingly but without an intention to A, (ii) A-ing is bad, but (iii) the agent is not blamable for A-ing. The test case is this: An executive implements a new program which she knows significantly increases sales in Massachusetts but decreases by much less sales in New Jersey. Decreasing sales in New Jersey is bad (it is said), but not something for which the executive is blamable. And "yet, people generally say that the executive intentionally decreased sales in New Jersey" (2004, 220).

3. METHODOLOGICAL CONSIDERATIONS

The first point to make is that, as we saw in section 1, responses to surveys about scenarios used in thought experiments are not *ipso facto* intuitions, that is, they are not *ipso facto* judgments which express solely the subject's competence in the deployment of the concepts involved in them in response to the scenario. The responses that people make in such circumstances are apt to be the product of a number of different factors, among which are how they understand the task, their background beliefs, empirical and nonempirical, how they think what they say will be taken, loose analogies they may draw with other sorts of situations, how they understand the scenario, whether they pay adequate attention to relevant details, whether they think clearly and hard enough to see what to say in response to the kind of question asked, assuming they understand it correctly, how they think that their interlocutor will (or interlocutors generally would) understand what they say or more generally what they would be trying to convey by what they say or how they respond, as well as perhaps various shortcuts or rules of thumb in reasoning, or plain mistakes.

The very fact that in both the examples we considered there is variation in the subjects' responses shows that, on the assumption that all the respondents are fully competent speakers of English and that there is no ambiguity in the terms used to describe the scenarios and in the questions posed, the responses are not

data in the form of intuitions about the application conditions of the concepts involved in their judgments in response to the described scenarios. For identical competencies would yield identical judgments in response were the responses made solely on the basis of those competencies and identical understandings of the scenario, task, question, and adequate thought.

The task when presented with responses which we know are not (at least all) intuitions is to try to factor out the contribution of competencies from the other factors. This requires an understanding of what the various factors are that may influence responses and enough information about each subject to be able to say with some confidence what factors are at work. It is clear that in the circumstances in which these surveys are conducted we do not have this kind of information.

To this it may be responded that we don't need it because we may assume that how most people will respond indicates what sort of response is most likely to express the judgment that would be appropriately made solely on the basis of the concepts involved in response to the scenario and task as they are intended to be understood. However, there is clearly no *a priori* reason to think that this is so. In other contexts, it is clear that many people do systematically give responses which are not expressions of their full competence in the application conditions of the concepts which they deploy in their judgments in response to the situation as described.

The gambler's fallacy provides a good example. Even if an experimenter describes a coin as fair, many subjects will judge it becomes more likely that it will come up tails after a run of heads. Since we think that those who so judge are in error, we must suppose that their responses are not in fact an appropriate expression of their underlying competence in the concepts involved in response to the question and scenario understood as intended, that is, their responses do not express intuitions about the scenario as it is intended to be understood.

For another example, imagine the following probe: Mr. Smith, a normal fit man in his twenties, is standing by the window in his living room. He hears the phone ring on the side table by the couch on the other side of the room. He walks to the side table across the open floor between him and the phone and picks it up. Which of the following is true: (A) Mr. Smith tried to walk across the living room; (B) Mr. Smith did not try to walk across the living room; or (C) Neither (A) nor (B). It is plausible that many people would choose (C), as did quite a number of philosophers in the 1950s. Attention to the circumstances in which people do readily accept that someone has tried to do something offer an explanation: Someone tries to do something only when there is or might have been some difficulty involved or significant prospect of failure. It is now widely accepted that this is mistaken and that (A) rather than (B) or (C) is the correct response. Why? As Grice and O'Shaughnessy have pointed out (Grice 1989, 17; O'Shaughnessy 1973, 387), ordinary responses are sensitive not so much to what the objective facts about the case are as what the speaker thinks they are. If a speaker has good reason to believe that someone faces a difficulty in some task he is engaged in (starting a car, for example) of a sort which makes it very unlikely that he will succeed, he (the speaker) does not know that the agent will succeed: but as O'Shaughnessy notes, he at least knows that the agent in question is trying. Yet this is compatible with the

agent himself knowing full well that there is no difficulty about the matter at all. If we remove the doubts of the speaker or if he had not existed at all, the facts of the case would remain the same. As O'Shaughnessy points out, this device for stripping away confidence in success to leave knowledge of trying could be applied in any case of intentional action, which leads to the conclusion that trying is present in all intentional action. In this case, seeing that a mistake has been made in choosing (C) is possible without immediately having a diagnosis of what the mistake is.[23] Moreover, in this case, once we have worked through it, and returned to the original question, we can see that the right answer is (A) after all, just as when we work through it we can see the mistake in the gambler's fallacy.

The situation may be compared with one in which we are trying to test a hypothesis about the gravitational attraction which the earth exerts on objects close to its surface by allowing a disk to slide down a variety of different inclined planes (our different subjects) and measuring the time it takes to reach the bottom (eliciting responses to our question about the scenario) when we do not know anything about the angle or length of incline, or the amount of friction between the disks and the different inclined planes (the various factors which can influence responses). In this case, computing a mean and standard deviation from our measurements gives us a number, but not a number which can be used to say very much about what it was that we were interested in testing. There would be an illusion of having produced an objective measurement relevant to a question—after all, we have a set of data points from which we can extract a mean and a standard deviation—but the results would be relatively worthless. Suppose we were to carry out our experiment in two different locations, using local inclined planes, say, at Hong Kong University and at Rutgers University, and found that we got different

23. A diagnosis is provided by Grice's theory of conversational implicature. If whenever (or typically) someone A's intentionally he tries to A, then asserting that someone A-ed, in conditions in which it is clear that a description of the action is chosen under which it was intentional, carries the information that he tried to A. However, since one can evidently try to A without succeeding, if one asserts that someone tried to A when he is known to be in a position to be more informative, and he otherwise can be assumed to be obeying the Cooperative Principle, his interlocutor will suppose that he intends to convey something more than what he says and that in asserting something that would be true even if an attempt to do something failed he is plausibly taken to be suggesting that the agent will or has failed or that at least the difficulty he faces is greater than might normally have been expected or anticipated. Thus, when confronted with the scenario in the text, we do not want to say (A) is correct because this would standardly carry a false conversational implicature; we don't want to say (B) is correct because this would entail something false, that the agent did not intentionally walk across the room. We are then left with (C) as the least problematic response but distinctly uncomfortable response. Adams's response to Knobe's survey results is patterned after this sort of explanation. As we saw, Knobe suggests that the reasoning involved, particularly in the case of the second iteration of the exchange, is too complex for people to go through in a few seconds. But the reasoning underlying our grasp of just about any conversational implicature looks forbiddingly complex if we write out a rational reconstruction of it! See Grice's articulation of the reasoning involved in working out the implicature of someone's saying of a new bank employee, "he hasn't been to prison yet" (Grice 1989, 31), or the accounts Searle gives of the reasoning underlying grasp of indirect speech acts (Searle 1979). Figuring out what people intend is, not surprisingly, a skill we have developed to a very high degree of proficiency.

means at each location using the sets of locally selected inclined planes! What would it show?

Part of the point of classifying intuitions on the basis of their etiology is that it helps to bring out more clearly some of the methodological difficulties that arise in conducting thought experiments. The initial data we get in the form of responses cannot be assumed without further ado to be data points of the sort we want for the purposes of analysis. The conflation of "response to a scenario" with "data point about the analysis of a concept" is like the conflation of "time interval" and "data point about the gravitational attraction the earth exerts on objects close to its surface." The former amounts to the latter only relative to a host of assumptions about the experimental conditions.

Do we have any reason, however, to think that in the cases in question there are any systematically misleading factors at work? Let us suppose that the results of the surveys are robust in both our examples. That is, in studies involving the proper name probes, if repeated using the same probes or relevantly similar probes, we would find very similar results across subpopulations of East Asians and Westerners, and, in the case of probes involving responses to questions about acting intentionally in the help/harm scenarios and similar scenarios, we find similar patterns of response. It might be said that where there is no particular reason to think that there are factors that are at work which would have the result that most responses failed to line up with the judgment that would correctly be made solely on the basis of competence in the concepts involved, methodologically the proper thing to do is to take them at face value.

As the gambler's fallacy and similar examples show, this would not be advice that in general leads to the right result. Furthermore, in the case of our cross-cultural surveys about proper names, the results of the surveys themselves suggest that this is not a sound response. For in the different subgroups, different responses, at least to the Gödel case, predominate. If these differences indeed do have to do with cultural differences between members of the groups, as is the hypothesis, then we have good reason to think that in neither case can we take how most people respond to indicate what the underlying competence in the use of the concepts dictates.

This might be thought, however, just to be more grist for the mill of the skeptic about the traditional first-person-oriented method of analysis. Surely what the variation within and across cultures shows is that the method of thought experiments is a very uncertain business and that what we need is even more empirical research and even less armchair research.

To see what is wrong with this thought and to highlight a fundamental difficulty with analysis by surveys, let us shift our attention for a moment to another *a priori* field of inquiry. Consider the following proposal for how to decide whether mathematicians have got the "folk concept"[24] of number right. Let us suppose we conduct a survey in a park in New York.

24. The use of the slightly derogatory term is no accident in these discussions. It is what in the nature of what used to be called a persuasive definition.

Consider the following two number series:
(a) 0, 1, 2, 3, . . .
(b) 1, 3, 5, 7, . . .
Which of the following claims is correct?
(A) (a) has more members than (b).
(B) (b) has more members than (a).
(C) (a) and (b) have the same number of members.

While I have not conducted this experiment, I speculate that a preponderance of the responses will indicate (A) as the correct answer. Let us imagine now conducting a "cross-cultural" survey with this probe, say with a group of undergraduates at MIT and a group of undergraduates at the School of the Fine Arts in New York. I boldly hypothesize that in the first group one would find that a preponderance of respondents would choose (C) while in the second a preponderance of respondents would choose (A).

In this case, provided that we have had a certain amount of training in mathematics, it is easy to pick out the correct answer, (C), and to explain why it is the correct answer. It is also relatively easy to explain to someone who gets it wrong by choosing (A) wherein his mistake lies—though to be sure some people will find it easier to grasp than others, and some may never get it. Of course, figuring out how to think about the cardinality of infinite series was not a simple or easy matter, even if once grasped it is not difficult to understand. Many conceptual advances have this character. For someone with appropriate training, answering such questions is relatively easy, even if it is the first time that he encounters it. For someone coming to it fresh and reflecting on these things for the first time, it can be anything but easy.

I don't suppose anyone will think that it is out of the question that the survey results should turn out as I have predicated. And I don't suppose that anyone will think that if they did, it would show that there was some fundamental difficulty with the procedures of the *a priori* discipline of mathematics. I don't suppose that anyone would think that the first imagined survey would show that mathematicians were not operating with the same concept of number (specifically natural number) as ordinary folk, or that the second imagined survey would show that there was some deep problem with the assumption of the universality of the number theoretic intuitions of mathematicians. But why not?

The answer in this case is relatively clear. It takes some thinking and reflection to see what the right answer should be on the basis of the concepts which are involved. We need to think carefully about what is involved in two series having the same number of members and recognize that in the case we are confronted with the series have an infinite number of members. Working out that in the finite case a one-to-one correlation suffices for equinumerosity, we can see that projecting this to the infinite case leads to the conclusion that the odd numbers are equinumerous with the naturals. In other words, a quick judgment is not called for, but rather a considered reflective judgment in which the basis for a correct answer is revealed.

It is clear that two sorts of training are relevant to being able to see what the right thing to say is in this case. The first is general analytical training relevant

to thinking correctly and clearly about problems in structured ruled-governed domains. The second is training in mathematics which involves familiarization with working with number systems. This is not a matter of acquiring new concepts! It is a matter of gaining greater sensitivity to the structure of the concepts through reflective exercises with problems involving those concepts. In other words, training and expertise matter to how well positioned people are to give correct answers to questions like the one posed above.

Similarly, people can train themselves to see what is around them more accurately and in more detail, as do artists through the exercise of drawing or painting from a scene. People can train themselves likewise to be more attentive to changes in scenes which they would otherwise be apt to overlook. It is clear that in such cases the observations of trained observers are more valuable than those of untrained observers.

About some matters people can make confident and correct judgments without special expertise or training. In other areas, training and expertise significantly improve their ability to make correct judgments. What is the case with respect to intuitions we make in response to scenarios in conducting thought experiments? What we can expect is that in some cases nearly everyone competent with the use of the concepts in question can make confident and correct judgments. Anyone competent with respect to the basic numerical concepts knows that 2 is greater than 1, for example, and anyone competent in the use of the concept of agency knows that if someone is blown off a cliff by the wind it is not something he did at all and so not something he did intentionally. But we should not expect that in every case in which we are called on to make a judgment we are at the outset equipped to make correct judgments without much reflection. Our concepts generally have places in a family of related concepts, and these families of concepts will have places in larger families of concepts. How to think correctly about some cases we are presented can be a matter that requires considerable reflection. When a concept, like that of justification, is interconnected without our thinking in a wide variety of domains, it becomes an extremely complex matter to map out the conceptual connections and at the same time sidestep all the confusing factors.

Let us return to the thought that the difficulties in conducting thought experiments and the variations in the responses which are produced when ordinary people are quizzed about scenarios used in thought experiments by philosophers show that the whole enterprise of conceptual analysis as traditionally conceived is in jeopardy and that more rather than less empirical research is needed to see how to proceed. First of all, it should be clear from the imagined scenario involving the relative cardinality of the series of naturals and odd numbers that the conclusion does not follow from the facts about variation in response to surveys. What is indicated rather is that we have an area in which the complexities presented to subjects make it difficult to see clearly what one should say. What is called for then is the introduction of training and expertise in sorting out the various confusing factors that may be at work. Using surveys of untrained people to settle issues where there are conceptual knots in our thinking is fundamentally methodologically misguided.

We should not expect antecedently that untrained subjects should be in an especially good position to give judgments in response to scenarios involving difficult questions about the semantics of proper names, for this is a domain of considerable complexity where our ordinary vocabulary is not especially precise. We should instead expect that the relevant experts in the field of philosophical semantics will be better placed to give answers which focus on the right features of the cases and what they are supposed to be responding to.

This general point can be illustrated by attending to a problematic feature of the probes used in the surveys about proper names reported above, which is how the question posed to the subjects is worded: "When John uses the name "Gödel," is he talking about . . . " For anyone at all familiar with work in the philosophy of language, it is immediately evident that the question does not clearly distinguish between two things: whom John intends to be talking about (or speaker's reference) and who the name John uses refers to, taken literally in the language he intends to be speaking (semantic reference). Experts may well negotiate this infelicity in the formulation of the question without much difficulty, but that is because they have some relevant expertise about hard-won distinctions developed in the field and will likely understand what is intended. This highlights the importance of the point in section 1 that it is desirable that the subject of the thought experiment have expertise in the relevant field.

Indeed, once the distinction has been introduced, it can be seen that there is only one correct response to the question understood as intended, given the description of the situation in the probe (and what must be taken as an implicit assumption of it for it to so much as make sense, namely, that John is told what he is told by someone who uses the name "Gödel" in telling him about Gödel). In the description of what John was told, the name Gödel is used to refer not to the discoverer of the incompleteness of arithmetic, but to the person who stole the manuscript. So what John is told is something that is about Gödel and not about Schmidt. So when John reports what he was told, he reports something about Gödel, for what he was told he was told in English and "Gödel" as used in that sentence referred to Gödel, and when he says, for example, "Gödel proved the incompleteness of arithmetic," he intends to be reporting what he was told using English, and consequently, insofar as the question is about what "Gödel" refers to in his uses of it relative to the language he is speaking, it refers to Gödel. That is, the way the thought experiment is set up, there is only one answer that is acceptable. It is not described in a way that allows any other correct response.[25] This highlights, as noted in section 1, the importance of expertise in the experimenter in relation to the subject matter as well as the method of thought experiments to developing useful scenarios and questions about them.

What is called for is the development of a discipline in which general expertise in the conduct of thought experiments is inculcated and in which expertise in different fields of conceptual inquiry is developed and refined. There is such a discipline. It is called philosophy. Philosophers are best suited by training and expertise to conduct thought experiments in their areas of expertise and to sort out

25. Kripke's own discussion artfully avoids the difficulty on display here.

the methodological and conceptual issues that arise in trying to get clear about the complex structure of concepts with which we confront the world. A lot of the problems we confront are very difficult, and so it is not surprising that in many areas there is lively debate, but it would be a mistake to think that the way to resolve such debates is to return to questioning untutored subjects in just the places where there is evidence of the most difficulty in coming to a clear view.

It might be suggested that we could ameliorate the problems of interpreting what untutored subjects say about scenarios in thought experiments by controlling for their understanding of what they are being asked to do and their skill in responding correctly. We can give them explicit instruction in the method of thought experiments, and we can give them some examples which will test to see whether they can make distinctions relevant to responding just to the scenario on the basis of their competence in the use of the relevant concepts, especially with respect to cases in which we know people are apt to fall into errors from one or another known factor. We can then select the subjects who are best at responding on the basis of competence in the use of concepts and present the scenarios we are interested in to them. Furthermore, we can fill in for subjects some of the relevant background in the conceptual field intuitions about which we are interested in so that they can bring to bear a sophisticated understanding of what the issues are, much as in training students in mathematics we put them in a better position to respond to questions in number theory on the basis of their competence in the deployment of the concepts in question. It seems clear that this would significantly improve our confidence that the answers we got in response to surveys would reflect intuitions *per se* about the scenarios. But the logical end point of this process is to give the subjects we want to run our tests on an education in philosophy, and it is to admit that training in philosophy puts one in a better position in general to sort out what the proper response is to a scenario in a thought experiment.

There is a further general point which it is important to make. While it is not unreasonable to think that some empirical research on the foibles people are apt to fall into in thinking may shed light on errors that ordinary people (and even philosophers) make in response to thought experiments (e.g., Horowitz 1998), to sort out ultimately where mistakes are being made rather than the contours of our concepts delimited, those conducting the experiments must themselves have a good enough reflective grasp on the correct application of the concepts concerned that they can describe correctly what is going on. One cannot diagnose an error unless one knows that an error has been made. This is not something that experimenters will be able to uncover by asking untrained subjects. Rather, it is something that they must know themselves. They must know from the first person perspective and that means that they must (a) be able to respond correctly to the scenarios on the basis of their competence and (b) see that they have done so. This is clearly at work in the diagnosis of the mistakes that many people appear to make in probabilistic reasoning in certain circumstances. We know that the gambler's fallacy is a fallacy despite the evidence that many ordinary people make it because we see that it follows from each trial being independent of the others that the probability of a fair coin coming up heads (tails) is one-half each time it is flipped.

Let's take stock by reviewing the various charges that have been made against the procedures of philosophers in Machery et al. (2004) and the responses to them. After that, I will conclude by considering in what respects surveys can be an aid to philosophy and the relative merits of the third versus first person approach.

(a) Semantic intuitions are culturally variable, and, indeed, interpersonally variable, and perhaps variable within a given person across times. Therefore, the assumption of philosophers of the universality of semantic intuitions is "just spectacularly misguided."

The surveys cited show nothing of the sort, and they could not do so. It is a mistake to label the responses to the surveys "semantic intuitions," as we have seen. To call them semantic intuitions suggests that the responses are being guided by an underlying competence in the relevant concepts. And this cannot be so on the assumption of the experiments that the subjects in question share the same concepts, for this is to assume that they have identical competencies. All the surveys show is that philosophically untutored subjects do not all give the same responses to the scenarios involving the reference of proper names and that there are different tendencies in the answers between Hong Kong undergraduates and Rutgers undergraduates. Philosophers have certainly never assumed the universality of responses to thought experiments about proper names or anything else! Once the relevant distinctions are on the table, it can be seen that what is spectacularly misguided is the assumption that responses to the relevant scenarios are *ipso facto* expressions of semantic intuitions. (In fact, the experimenters adhere to this assumption only when it suits their purpose, for they do not get significant differences between groups in the Jonah cases, and consequently they do not take those to be showing anything about semantic intuitions. If they did, they would have to say that the semantic intuitions among the East Asian participants were inconsistent.)

(b) It is "*wildly* implausible that the semantic intuitions of the narrow cross-section of humanity who are Western academic philosophers are a more reliable indicator of the correct theory of reference . . . than the different semantic intuitions of other cultural or linguistic groups."

When we have distinguished between responses to scenarios and semantic intuitions, it is clear that it would indeed be wildly implausible to think that the semantic intuitions of philosophers were more reliable than those of others. Sameness of underlying competence ensures sameness of semantic intuitions. It is *not*, however, wildly implausible to assume that philosophers' responses to scenarios are more likely to be reliable indicators of the correct intuitions. Expertise both in the relevant fields and in the methodology and conduct of thought experiments is clearly relevant to reliability (as we have seen). It is so in other fields and it would be risible to suggest that philosophy is an academic discipline which unlike every other fails to develop any subject matter or methodological expertise in its practitioners. The thought expressed in this objection seems to be based on the conflation of responses to scenarios in thought experiments with genuine intuitions, in which case the variation in responses allows one to read off a variation in semantic intuitions! If semantic intuitions express directly competence in the deployment of

the relevant concepts, we are driven immediately to a bizarre cultural relativity of the application conditions of concepts. But this is incoherent. There might be differences in what people mean by the same words or the semantic properties they attach to words between different linguistic groups, but there cannot be a difference in what meanings mean between different linguistic groups. That is a category mistake. If people mean the same by the words they use, they express the same concepts with them and they and the words that express them have the same application conditions.

(c) Theory bias: "given the intense training and selection that undergraduate and graduate students in philosophy have to go through, there is good reason to suspect that the alleged reflective intuitions may be reinforced intuitions."

Imagine someone saying, "given the intense training and selection that undergraduate and graduate students in mathematics go through, there is good reason to suspect that the alleged reflective intuitions they have about mathematical matters may be reinforced intuitions. In the absence of a principled argument why mathematicians' intuitions are superior to untutored subjects, their project smacks of narcissism in the extreme." The charge is clearly wild! Intense training and selection goes on in every academic discipline. There is always the worry that in the course of training one will become entrenched in a received view and resistant to evidence that it is false, but this is no reason for a blanket condemnation of a discipline or a reason to think that training and expertise are worthless. Furthermore, it is itself clearly an empirical charge, which would require empirical evidence in support, and in the case of philosophy, that would involve as a component establishing what were the correct conceptual accounts in areas that philosophers have dealt with, and so employing the very procedures that philosophers employ while somehow avoiding their supposed mistakes. This presupposes that whatever problems there are, they are correctable with appropriate reflective attention.

(d) "Since the intuitions philosophers pronounce from their armchairs are likely to be a product of their own culture and their academic training, in order to determine the implicit theories that underlie the use of names across cultures, philosophers need to get out of their armchairs." The ground for the conclusion reached here has been dealt with, but the conclusion drawn embodies another mistake, namely, that an implicit theory underlies the use of names. When we learn a language, we do not learn a semantic theory—we learn how to use words in various categories. We acquire a complex skill which is manifested in our use of the words we learn and in the judgments we make using them. Those judgments are a guide to the proper analysis of the concepts we express with them and of the semantics of the expressions we use. The competencies we acquire determine the semantic properties of the words we use and do not constitute an empirical theory about them.

No rational skepticism about the traditional methods of philosophy can be grounded on the results of surveys of untutored subjects. The next question to address is whether surveys of untutored subjects offer any advantage over the traditional approach.

We have seen that there is no reason to think that responses of untutored subjects are *ipso facto* more reliable than those of professional philosophers who

have considerable training in their subject fields. On the contrary, there is reason to think that professional philosophers are generally more reliable. This is not to say that professional philosophers always get it right, or there would be much less dispute in philosophy than there is. But this is a reflection of the difficulty of the field rather than the irrelevance of training, and just tends to cast more doubt on the strategy of querying the untutored. Should we then just throw the surveys out the window?

Not so fast! First, plausibly, surveys have a useful role to play as a corrective to two connected vices that can develop in philosophy, theory bias—letting one's responses be guided by one's theory rather than by one's competence in the deployment of the relevant concepts—and what we might call theoretical insularity—failing to interact enough with and respond to different or opposing views or problematic data. Consulting others who have not worked their way into a particular picture of a problem area is valuable and, in general, a certain amount of friction in developing a view or position is helpful in thinking it through. Philosophers are well aware of this and philosophy is not, as I noted in section 1, just a solitary affair but a community effort. In other philosophers we find a sophisticated audience with respect to which to test accounts in various problem fields. Even so, contact with responses to thought experiments of completely untutored subjects can often be useful, and it is in this connection that I think "experimental philosophy" plays a constructive role, for it supplies a systematic source of data on ordinary untutored responses to scenarios used in thought experiments. If my experience is a guide, philosophers do this informally in teaching and in conversation with nonphilosophers. Experimental philosophy is a more formal approach to this ongoing practice. It provides a check on one's assumptions about the conformity of one's own responses to those of others and especially those who have not been trained in philosophy. This yields occasional surprises which are valuable even if one decides that they can be explained away, for they have the potential for revealing unrecognized conceptual connections which play a role in even incorrect responses and for providing greater insight into methodological difficulties in conducting thought experiments, which can be brought back to inform one's understanding of one's own responses to thought experiments. (Something similar can be said more generally of psychological studies of ordinary reasoning, even if one occasionally thinks their interpretation would itself benefit from great conceptual clarity.)

Second, surveys provide a method of testing diagnoses specifically of survey results, though as the discussion below of the survey involving an executive's introducing a new sales program shows, this is not always a straightforward matter.

Third, surveys also provide evidence about strength of tendencies in competent populations to certain responses across a variety of scenarios in which relevant factors are systematically varied which it is hard to get otherwise, and this additional information arguably may point to patterns of connections between concepts and sources of diagnoses which it would be otherwise hard to identify (see, for example, the data in Nadelhoffer 2006).

Nonetheless, a resolutely third person approach, especially by survey, presents some special difficulties, and we must ultimately return to the first person approach to achieve the goals we set ourselves in philosophy.

When the roles of experimenter and subject are distinct, the experimenter is faced with some difficulties he does not face when they are the same. First, he is faced with the task of determining that the subject understands the task, namely, to respond solely on the basis of competence in the use of the relevant concepts. Second, he is faced with the task of describing the scenario and question in a way that wards off misunderstandings and tacit assumptions about it which may result in responses to a situation type other than the one the experimenter wants a response to. This is less difficult with trained subjects, but much more difficult with untrained subjects. Clearly, if the roles of experimenter and subject are collapsed, these problems are significantly reduced. The experimenter is in the best position to know how the scenario is intended to be understood, what the task is, and how the question is to be understood, and when the experimenter is the subject, the subject is in as a good a position to respond appropriately, with respect to these factors, as he can be. In this respect, the first person approach has an advantage.

The survey approach has another disadvantage, which is that often in testing a hypothesis about a conceptual connection it is useful to run through a number of different scenarios, some developed on the basis of how responses go to previous scenarios and explanations offered of them. We see this in fact in the discussion of the Simple View reviewed above. If we are conducting the thought experiments in the first person, this is a relatively rapid process. Indeed, it is a relatively rapid process if we are conducting it in a conversation with a reasonably sophisticated interlocutor. But it is not a relatively rapid process if we are conducting it with surveys of ordinary people, and in this connection one might wonder whether in the social science approach we have hit on the most efficient way of making progress and guarding against the dangers of theory bias and theoretical insularity.

In this connection, consider the survey reported in the last paragraph of section 2. An executive adopts a new program that decreases sales in New Jersey but increases them in Massachusetts and overall. The decrease in New Jersey is bad (or said to be) but the executive is not blamable for it. Most subjects are reported as saying that the executive decreased sales in New Jersey intentionally. The scenario, though, leaves something unspecified, namely, whether the way the increase in sales in Massachusetts was achieved involved as a means something that must decrease sales in New Jersey, such as transferring salespeople or advertising budgets from the one area to the other. In this case, the decrease in sales in New Jersey was likely conceived of as part of the plan for increasing sales overall. Unless minimally we rule out that the subjects who respond by saying the executive decreased sales intentionally because they assumed the means to his end involved decreasing sales in New Jersey, we don't know enough to see whether the responses bear on the issue. Now, if we philosophers can reasonably trust our own intuitions or at least those of our so far uncommitted colleagues, with a good interlocutor this difficulty can be uncovered quickly and then relatively quickly a new scenario against which to test our reactions can be described in enough detail to make clear that the decrease in sales in the one area is not to be conceived of as a means to the end of increasing the sales in the other area.

Finally, though, whatever the merits of the survey method, we must bring our subject back to the first person perspective. For our object is to learn about the world that we use our concepts in thinking about.[26] What we want is not knowledge of what our words mean or even second-order knowledge of our concepts. We seek knowledge of the contours of and connections between things which are reflected in the contours of and connections between our concepts of them. We want conceptual knowledge in the sense of knowledge of truths whose ground is our conceptual competence, and which employ rather than mention the concepts of interest. We do not want to know of a conceptual truth that it is true on the basis of empirical research. We do not want to know even just that it is true on the basis of the contained concepts, that is, that it is a conceptual truth. For, first, we could know that without knowing the conceptual truth. And, second, even if we did know it on that basis by understanding it and deducing it from a second-order statement, we would not thereby know it in the way we wish to know it. This would be like

26. At least some practitioners of experimental philosophy (some of the time) may have a different goal in mind. If so, the bearing of their enterprise on conceptual analysis in the sense in which philosophers have been interested in it becomes less clear. For example, Knobe and Burra take Alfred Mele to task to describing the object of conceptual analysis as to produce a material mode biconditional. They say (Knobe and Burra 2006a, 332), "The problem with such an account is that it seems to say nothing about people's *concepts*. (It would tell us, not about people's concepts, but about the actual properties in the world that these concepts pick out.) We take it that a person's concept is a particular type of mental representation. Hence a theory about people's concepts must be a theory about particular mental representations that people possess." They say later, "the usual reason for investigating folk concepts has nothing to do with learning about which properties these concepts pick out. . . . We study these folk concepts because we are interested in questions about how people ordinarily come to grips with certain aspects of their environments" (Knobe and Burra 2006a, 338). There are a number of issues that need to be sorted out here, and there will be room here only for the beginnings of a proper discussion. First, Mele is right about what the traditional object of conceptual analysis is: to arrive at a conceptual truth about what we apply our concepts to. Second, in any case, it should be noted that if a material biconditional expresses a conceptual truth, then while it is not about the concepts involved, knowing that it is a conceptual truth gives us information about the competence that is involved in grasp of the concepts involved, namely, that it must endorse the biconditional (not of course that it be obvious to anyone who possesses the concept). Third, however, talk of mental representations puts a different cast on the whole discussion. There are a number of different ways of understanding what this comes to, some more and some less abstract. Some of them clearly are themselves parts of thoroughly empirical theories of cognition (à la the Language of Thought hypothesis). If this sort of thing is what Knobe and Burra have in mind, then the relevance of their enterprise to conceptual analysis as it is understood traditionally in philosophy would need to be worked out. For in that case "mental representation" is a theoretical concept, and it would need to be shown how its theoretical role was connected with our competencies in the deployment of concepts in the traditional sense. Certainly, our possessing and deploying in thought the concepts we do in thinking about the world and ourselves is not at all hostage to the correctness of any such empirical theory, for it is an open empirical question whether there are mental representations in the sense specified by the theory, but it is not an open empirical question whether we have concepts and deploy them in belief, desire, intention, and so on. We will not settle any questions about mental representations so understood from the armchair, but that is not what we were trying to do. At the least some more clarity is needed on this issue, since it appears that those in the experimental philosophy camp often think of themselves as working on traditional philosophical problems or making contributions to them, and they are so taken by many not in their camp. If the camps have been talking past each other, it would be good to know.

knowing the Pythagorean Theorem on the basis of authority and knowing that it had an *a priori* proof, without understanding why it was true *a priori*. In mathematics, the object is to come to know mathematics truths on the basis of understanding. And similarly in philosophy, in conceptual investigations, the goal is to come to know conceptual truths on the basis of understanding. This means that each of us must finally come to know the statements which express conceptual truths in the areas we are interested in, and when we employ the method of thought experiments in pursuit of this, this means that eventually we must cease to rely on others' responses and see for ourselves what is so about the scenarios with respect to the relevant concepts on the basis of our grasp of them. We must decide whether most ordinary people get it right or not for ourselves on the basis of our grasp of the relevant concepts, and we will not have settled the issue we are concerned with until we have done so. The first person perspective in thought experiments is therefore methodologically primary, as it is the only perspective from which we can attain the goals of philosophical inquiry.

4. CONCLUSION

The survey method of conducting philosophical thought experiments is not without its value but faces considerable methodological difficulties and does not represent a method superior to the traditional methods of philosophy. It cannot stand alone if its goals are those of the traditional approach but instead must eventually lead us back to the first person conduct of thought experiments, as it is only from the first person perspective that knowledge of conceptual truths on the basis of understanding can be achieved.

REFERENCES

Adams, Fred. 1986. "Intention and Intentional Action: The Simple View." *Mind and Language* 1: 281–301.

Adams, Fred, and Steadman, Annie. 2004a. "Intentional Action and Moral Considerations: Still Pragmatic." *Analysis* 64(3): 268–76.

———. 2004b. "Intentional Action in Ordinary Language: Core Concept or Pragmatic Understanding?" *Analysis* 64(2): 173–81.

Anscombe, G. E. M. 1957. *Intention*. Oxford: Blackwell.

Bealer, George. 1998. "Intuition and the Autonomy of Philosophy." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 201–39. Lanham, MD: Rowman & Littlefield.

———. 2002. "Modal Epistemology and the Rationalist Renaissance." In *Conceivability and Possibility*, ed. T. Gendler and J. Hawthorne. Oxford: Clarendon Press.

Beaney, Michael, ed. 1997. *The Frege Reader*. Oxford: Blackwell.

Bratman, Michael. 1984. "Two Faces of Intention." *Philosophical Review* 93: 375–406.

Burge, Tyler. 1979. "Individualism and the Mental." In *Midwest Studies in Philosophy: Studies in Metaphysics*, ed. Peter A. French, Theodore E. Vehling and Howard K. Wettstein, 73–121. Minneapolis: University of Minnesota Press.

———. 1987. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.

Cummins, Robert. 1998. "Reflection on Reflective Equilibrium." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 113–27. Lanham, MD: Rowman & Littlefield.

Foley, Richard. 1998. "Rationality and Intellectual Self-Trust." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 241–56. Lanham, MD: Rowman & Littlefield.

Goldman, Alvin, and Pust, Joel. 1998. "Philosophical Theory and Intuitional Evidence." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 179–97. Lanham, MD: Rowman & Littlefield.

Gopnik, Alison, and Schwitzgebel, Eric. 1998. "Whose Concepts Are They, Anyway? The Role of Philosophical Intuition in Empirical Psychology." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 75–94. Lanham, MD: Rowman & Littlefield.

Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Harman, Gilbert. 1976. "Practical Reasoning." *Review of Metaphysics* 29: 431–63.

Horgan, Terry, and Graham, George. 1998. "Southern Fundamentalism and the End of Philosophy." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 271–92. Lanham, MD: Rowman & Littlefield.

Horowitz, Tamara. 1998. "Philosophical Intuitions and Psychological Theory." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 143–60. Lanham, MD: Rowman & Littlefield.

Knobe, Joshua. 2003a. "Intentional Action and Side Effects in Ordinary Language." *Analysis* 63 (3): 190–94.

———. 2003b. "Intentional Action in Folk Psychology: An Experimental Investigation." *Philosophical Psychology* 16(2): 309–24.

———. 2004. "Intention, Intentional Action and Moral Considerations." *Analysis* 64(2): 181–87.

———. 2005. "Ordinary Ethical Reasoning and the Ideal of 'Being Yourself.'" *Philosophical Psychology* 18(3): 327–40.

———. 2006. "The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 130(2): 203–31.

Knobe, Joshua, and Burra, Aruda. 2006a. "Experimental Philosophy and Folk Concepts: Methodological Considerations." *Journal of Cognition and Culture* 6(1–2): 331–42.

———. 2006b. "The Folk Concepts of Intention and Intentional Action: A Cross-Cultural Study." *Journal of Cognition and Culture* 6(1–2): 113–32.

Kornblith, Hilary. 1998. "The Role of Intuition in Philosophical Inquiry: An Account with No Unnatural Ingredients." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 129–41. Lanham, MD: Rowman & Littlefield.

Kripke, Saul A. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Ludwig, Kirk. 1992. "Impossible Doings." *Philosophical Studies* 65(3): 257–81.

Ludwig, Kirk, and Ray, Greg. 2002. "Vagueness and the Sorites Paradox." *Philosophical Perspectives* 16: 419–61.

Machery, E., Mallon, R., Nichols, S., and Stich, S. 2004. "Semantics, Cross-Cultural Style." *Cognition* 92: B1–B12.

Malle, Bertram F., and Knobe, Joshua. 2001. "The Distinction between Desire and Intention: A Folk-Conceptual Analysis." In *Intentions and Intentionality: Foundations of Social Cognition*, ed. B. F. Malle, L. J. Moses, and D. A. Baldwin, 45–67. Cambridge, MA: MIT Press.

Malle, Bertram F., and Nelson, S. 2003. "Judging Mens Rea: The Tension between Folk Concepts and Legal Concepts of Intentionality." *Behavioral Sciences and the Law* 21: 563–80.

McCann, Hugh. 1991. "Settled Objectives and Rational Constraints." *American Philosophical Quarterly* 28: 25–36.

Mele, Alfred R., and Moser, Paul K. 1994. "Intentional Action." *Nous* 28: 39–68.

Nadelhoffer, Thomas. 2004a. "The Butler Problem Revisited." *Analysis* 64(3): 277–84.

———. 2004b. "On Implicit Testability and Philosophical Explanations." *Philosophical Writings* 27: 3–14.

———. 2005. "Skill, Luck, Control, and Intentional Action." *Philosophical Psychology* 18(3): 341–52.

———. 2006. "Desire, Foresight, Intentions, and Intentional Actions: Probing Folk Intuitions." *Journal of Cognition and Culture* 6(1–2): 133–57.

Naess, Arne. 1953. *Interpretation and Preciseness: A Contribution to the Theory of Communication*. Oslo: Jacob Dybwad.

Nahmias, Eddy, Morris, Stephen, and Nadelhoffer, Thomas. 2004. "The Phenomenology of Free Will." *Journal of Consciousness Studies* 11(7–8): 162–79.

———. 2005. "Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility." *Philosophical Psychology* 18(5): 561–84.

Nichols, Shaun. 2004a. "After Objectivity: An Empirical Study of Moral Judgment." *Philosophical Psychology* 17(1): 3–26.

———. 2004b. "The Folk Psychology of Free Will: Fits and Starts." *Mind and Language* 19(5): 473–502.

Nichols, Shaun, Stich, Stephen, and Weinberg, Jonathan M. 2003. "Meta-Skepticism: Meditations in Ethno-Epistemology." In *The Skeptics: Contemporary Essays*, ed. S. Luper, 227–47. Aldershot: Ashgate Publishing.

O'Shaughnessy, Brian. 1973. "Trying (as the Mental Pineal Gland)." *Journal of Philosophy* 70: 365–86.

Ramsey, William. 1998. "Prototypes and Conceptual Analysis." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 161–77. Lanham, MD: Rowman & Littlefield.

Searle, John R. 1979. "Indirect Speech Acts." In *Expression and Meaning*, 30–57. Cambridge: Cambridge University Press.

Shafir, Eldar. 1998. "Philosophical Intuitions and Cognitive Mechanism." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 59–73. Lanham, MD: Rowman & Littlefield.

Sosa, Ernest. 1998. "Minimal Intuition." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 257–69. Lanham MD: Rowman & Littlefield.

———. 2006. "Intuitions and Truth." In *Truth and Reason*, ed. P. Greenough and M. Lynch, 208–26. Oxford: Oxford University Press.

Stich, Stephen. 1998. "Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 95–1112. Lanham, MD: Rowman & Littlefield.

Wisniewski, Edward J. 1998. "The Psychology of Intuition." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, ed. M. DePaul and W. Ramsey, 45–58. Lanham, MD: Rowman & Littlefield.