

FLORIDA STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

THE RELEVANCE OF FOLK INTUITIONS TO PHILOSOPHICAL DEBATES

By

ADAM FELTZ

A Dissertation submitted to the
Department of Philosophy
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Summer Semester, 2008

The members of the Committee approve the Dissertation of Adam Feltz defended on May 19th, 2008.

Alfred R. Mele
Professor Directing Dissertation

Mark Isaac
Outside Committee Member

David McNaughton
Committee Member

Approved:

Piers Rawling, Chair, Department of Philosophy

Joseph Travis, Dean, College of Arts and Sciences

The Office of Graduate Studies has verified and approved the above named committee members.

For Mary Ann Feltz, in memoriam.

ACKNOWLEDGEMENTS

There are a number of people I would like to thank for helping me in one way or another to write this dissertation. I would like to thank my father for his support during my academic career. I would especially like to thank my brother, Norman Feltz, for keeping me in line and keeping things in perspective. I would also like to thank my committee—Alfred Mele, David McNaughton, and Mark Isaac—for helping me flesh out ideas through thoughtful feedback and insights. Moreover, I would like to thank Thomas Nadelhoffer who gave me insight, encouragement, and advice. I am thankful to Edward Cokely who offered invaluable advice on experimental design, statistical analyses, and implications of experiments. I am grateful to a number of very positive and fun graduate students here at Florida State who include, but are not limited to, Tracie Mahaffery, Chris Zarpentine, Clifford Sosis, Dan Deen, and Zach Martin. Last, but certainly not least, I would like to thank my lovely wife Silke. Without her constant support, motivation, and love I would not have been able to complete this dissertation (plus, she was an excellent editor!).

TABLE OF CONTENTS

| | |
|--|------|
| List of Tables | viii |
| Abstract | ix |
| | |
| 1. INTRODUCTION..... | 1 |
| 2. A PRIORI DEFENSES OF INTUITIONS..... | 9 |
| 1. Bealer's Approach..... | 10 |
| 1.1 The Conceptual and Empirical Case against Intuitions and the SJP..... | 11 |
| 1.2 Empirical Pressure on Bealer's View..... | 13 |
| 1.3 The Problem of Relativism..... | 13 |
| 1.4 The Problem of Infinite Intelligence..... | 17 |
| 2. The Mint of Nature Argument..... | 21 |
| 3. Sosa's Minimal Intuition..... | 24 |
| 3.1 Sosa's Argument for Minimal Intuition..... | 24 |
| 3.2 Some Problems with Sosa's Minimal Intuition..... | 26 |
| 4. Moderate Pessimism about Intuitions..... | 28 |
| 5. Conclusion..... | 32 |
| 3. EPISTEMOLOGY AND EMPIRICAL DATA..... | 34 |
| 1. Two Ways Epistemology Depends on Us..... | 35 |
| 2. Standard Analytic Epistemology and Empirical Data..... | 37 |
| 3. Describing Folk Intuitions..... | 39 |
| 4. Order Effects and Individual Differences..... | 47 |
| 5. The Relevance of the Empirical Data to Epistemological Debates..... | 50 |
| 5.1 Critical Implications..... | 50 |
| 5.2 Constructivist Implications..... | 51 |
| 6. Conclusion..... | 54 |
| 4. ETHICS, INTUITIONS, AND EMPIRICAL DATA..... | 55 |
| 1. Empirical Claims in Ethics..... | 56 |
| 1.1 Descriptions of Moral Platitudes..... | 56 |
| 1.2 Framing Effects on Moral Intuitions..... | 57 |
| 2. Ethics and Empirical Data..... | 60 |
| 2.1 Empirical Data and Objectivity..... | 60 |
| 2.2 Empirical Data and Practicality..... | 61 |
| 2.3 Worrisome Framing Effects..... | 61 |
| 3. More Evidence about Moral Intuitions..... | 63 |
| 3.1 Objectivity and Individual Differences..... | 63 |
| 3.2 Framing Effects and Individual Differences..... | 66 |
| 4. Implications of the Empirical Data..... | 68 |

| | |
|---|-----|
| 4.1 Critical Uses of the Empirical Data..... | 68 |
| 4.2 Constructivist Uses of the Empirical Data..... | 69 |
| 5. Conclusion..... | 72 |
| 5. FREE WILL AND EXPERIMENTAL PHILOSOPHY..... | 73 |
| 1. The Basic Positions, Intuitions, and Criticisms..... | 74 |
| 2. Additional Evidence and Intuitions' Stability..... | 79 |
| 3. Framing Effects, Reductionism, and Individual Differences..... | 86 |
| 4. Implications for the Free Will Debate..... | 89 |
| 4.1 Critical Uses of the Evidence..... | 89 |
| 4.2 Constructivist Uses of the Empirical Results..... | 91 |
| 5. Conclusion..... | 93 |
| 6. INTUITIONS AND ACTION THEORY..... | 94 |
| 1. The Knobe Effect..... | 95 |
| 2. Constructivist Explanations of the Knobe Effect..... | 96 |
| 2.1 Single Core Concept View..... | 96 |
| 2.2 Multiple Core Concepts Explanation..... | 96 |
| 3. Critical Explanations of the Knobe Effect..... | 98 |
| 3.1 Pragmatic Explanations..... | 98 |
| 3.2 Biasing Explanations..... | 98 |
| 3.3 The Attention Explanation..... | 100 |
| 4. Preliminary Glance at Folk Intuitions and Intentional Action..... | 100 |
| 5. More Data on Folk Intuitions about Intentional Action..... | 101 |
| 5.1 Order Effects and Sex Differences..... | 101 |
| 5.2 The Knobe Effect and Extraversion..... | 105 |
| 5.3 The Knobe Effect and Framing Effects..... | 106 |
| 6. The Relevance of Experimental Results for Traditional Action Theory..... | 108 |
| 6.1 Critical Implications..... | 108 |
| 6.2 Hopes for the Constructivist..... | 110 |
| 7. Concluding Remarks..... | 113 |
| 7. THE RELEVANCE OF FOLK INTUITIONS TO PHILOSOPHICAL DEBATES..... | 114 |
| APPENDICES | 120 |
| APPENDIX A. IRB Approval and Sample Verbal Consent Form..... | 120 |
| APPENDIX B. Stanley's Scenarios..... | 122 |
| APPENDIX C. Experiment with Stanley's Scenarios..... | 124 |
| APPENDIX D. Minimal and Attributer Cases..... | 126 |
| APPENDIX E. Simplified Cases..... | 128 |
| APPENDIX F. Bridge Cases..... | 129 |
| APPENDIX G. SAW's Scenarios..... | 130 |
| APPENDIX H. Ethics Experiment..... | 132 |

| | |
|---------------------------|-----|
| REFERENCES | 135 |
| BIOGRAPHICAL SKETCH | 143 |

LIST OF TABLES

| | |
|---|-----|
| Table 1: Folk Intuitions Predicted by Stanley..... | 40 |
| Table 2: Results of Stanley's Scenarios..... | 41 |
| Table 3: Results of the Minimal Cases..... | 43 |
| Table 4: Results of the Simplified Cases..... | 45 |
| Table 5: Results of the Bridge Scenarios..... | 47 |
| Table 6: Overall Means to SAW's Scenarios..... | 48 |
| Table 7: Men's Mean Responses to SAW's Scenarios..... | 49 |
| Table 8: Women's Mean Responses to SAW's Scenarios..... | 49 |
| Table 9: Mean Responses to Jim and Natives..... | 68 |
| Table 10: Responses to the Psychological Scenario..... | 87 |
| Table 11: Responses to Harm and Help by Order..... | 102 |
| Table 12: Men's Responses to Harm and Help by Order..... | 103 |
| Table 13: Women's Responses to Harm and Help by Order..... | 103 |
| Table 14: Overall Means to Harm and Help by Order..... | 104 |
| Table 15: Men's Responses to Harm and Help by Order..... | 104 |
| Table 16: Women's Responses to Harm and Help by Order..... | 105 |
| Table 17: Responses to Harm and Help by Postive and Negative Frame..... | 107 |

ABSTRACT

A large portion of philosophy done in the Western analytic tradition attempts to provide conceptual analyses which are tested by examples that elicit intuitions. These intuitions are, in turn, used as evidence either for or against a given analysis. In recent years, there has been much discussion of the uses of intuitions from empirically minded philosophers and psychologists. The basic strategy is to discover empirically how “normal” folks think about certain topics in philosophy. This application of folk intuitions to philosophy branches into roughly two basic approaches. The first is an attempt to show that in given domains, folk intuitions are not very reliable sources of evidence; hence, we have good reason to think that philosophers' intuitions are also not reliable sources of evidence in that domain. The second approach attempts to determine what folk concepts are. Once folk concepts are analyzed, they are then argued to be relevant to philosophical debates.

My guiding question for this dissertation is the following one: Why should philosophers care about folk intuitions? One answer is that we should want some philosophical analyses to be grounded in everyday concepts. I argue that there are presently no adequate a priori arguments for the reliability of philosophical intuitions in some philosophically relevant areas. Whether intuitions are reliable enough to ground philosophical analyses is an empirical question.

I review four domains where ordinary concepts have been argued to ground philosophical theorizing: (1) epistemology, (2) ethics, (3) free will, and (4) action theory. I argue that the available evidence suggests that we should be skeptical of intuitions in philosophy—but that skepticism does not entail radical skepticism. That is, the empirical studies reveal a wide variety of results which on the surface indicate that intuitions are not reliable (e.g., order effects or framing effects). However, I argue that these seeming instabilities are actually the results of stable differences in different groups of people. Hence, I argue that intuitions are stable in a surprising way—that different groups of people have stable intuitions. This intuition stability, while not the kind of monolithic stability many philosophers might desire, is argued to be sufficient to ground philosophical analyses in these domains.

CHAPTER 1: INTRODUCTION

Thought experiments play a vital function in analytic philosophy. With the use of thought experiments, philosophers think they can (a) refine conceptual analyses and (b) test theories. The ultimate goal of conceptual analyses is to provide a set of individually necessary and jointly sufficient conditions for a target concept. Thought experiments refine analyses because they provide tests of proposed necessary and sufficient conditions. That is, if one can find a counterexample to a proposed set of necessary and sufficient conditions, then the conditions provided as an attempted analysis of the concept are either not necessary or not sufficient. For example, let's say one attempts to give an analysis of concept C. One claims that conditions X, Y, and Z are jointly sufficient and individually necessary for C. If X, Y, and Z provide an adequate analysis of C, then no counterexample can be provided where (1) X, Y, and Z are present and we are not willing to say that C is exemplified, or (2) C is exemplified and not all of X, Y, and Z are present. If one can provide a scenario that can accomplish (1) or (2), then one has provided a counterexample to the attempted analysis of C.

For instance, take one of Gettier's (2000) counterexamples to the Justified True Belief analysis of knowledge. The Justified True Belief analysis of knowledge holds that the following conditions are individually necessary and jointly sufficient for a subject (S) to know a proposition P: (a) P is true; (b) S believes that P; and (c) S is justified in believing that P. Gettier claims that he can construct a case where it is obvious that (a)-(c) are present and we do not think the person in the example has knowledge. Suppose Smith has good evidence for the following proposition: Jones will get the job and Jones has ten coins in his pocket. Smith can then validly infer the following proposition: (d) the man who will get the job has ten coins in his pocket. However, Smith is the man who will get the job, and Smith has ten coins in his pocket, but Smith does not know either of these two facts. It "seems" clear that (a)-(c) are satisfied in this example: (d) is true; Smith believes (d); and Smith is justified in believing that (d). It is also clear that Smith does not know (d) because there is an unacceptable element of luck involved making (d) true (Gettier, 2000, p. 59). Hence, the Justified True Belief account is a failed analysis because in the example it "seems" (a)-(c) are satisfied and it "seems" to us that Smith does not know (d).

Roughly, these kinds of intellectual seemings are intuitions.¹ One has the intuition that (a)-(c) are met while at the same time one has the intuition that Smith does not know (d). Thus, intuitions are used to determine if a particular example is a counterexample to an analysis of a concept.

Intuitions are also used to test philosophical theories. One way intuitions have been used to test and construct philosophical theories is by reflective equilibrium. Reflective equilibrium is the method that counsels to the extent that theories allow inferences that are intuitively acceptable, those theories are correct. However, to the extent that theories license inferences that do not accord with our intuitions, those theories are to be rejected (Goldman, 1986). Thought experiments are often used in the process of reflective equilibrium. When one is presented with a thought experiment and the theory licenses an inference about that thought experiment that does not accord with some of our deeply held intuitions, then that theory has a *prima facie* strike against it. For example, some philosophers endorse the Principle of Alternative Possibilities, which asserts that one is morally responsible for an action only if one could have done otherwise. However, Frankfurt style cases generate intuitions that one can be responsible even if one could not have done otherwise. Because Frankfurt style cases generate intuitions that one can be morally responsible even if one could not have done otherwise, the intuitions generated by Frankfurt style cases count against the truth of the Principle of Alternative Possibilities. If a theory accords more or less with our intuitions, then those intuitions are thought to be evidence that the theory is true. Hence, our intuitions are used to test the truth or adequacy of our philosophical theories.

In recent years, psychologists and empirically minded philosophers have investigated the uses of intuitions. Their main goal is to discover systematically and empirically how “normal” folks think about philosophically relevant issues. These empirically minded theorists then use those folk intuitions in philosophical debates. This use of folk intuitions in philosophy branches into roughly two basic approaches. The first is an attempt to show that in some domains, folk intuitions are not very reliable sources of evidence. Because of their unreliability, some have claimed either (a) we cannot use folk intuitions as evidence for theories or conceptual analyses or

¹ It should be noted there is no consensus in philosophy about what intuitions are. Some think intuitions are simply beliefs (Lewis, 1983), some think they are inclinations to believe (van Inwagen, 1997), some think they are immediate seemings (Goldman & Pust, 1998), and for others they are seemings with some sort of special aura (Claxton, 1998). For a more extensive review, see Feltz and Bishop (in press). Following Feltz and Bishop, I give intuitions the following neutral characterization: intuitions are quickly formed judgments “of the sorts that have played vital roles in the development of philosophical theories over the past half century or so” (in press). By ‘quickly’, we do not mean judgments that are formed hastily but rather judgments that are formed immediately (which is consistent with them being formed after prolonged reflection).

(b) we have reason to think that philosophers' intuitions are also not reliable sources of evidence. The second approach attempts to determine what the folk intuitions are. Once the folk intuitions are discovered, they are then used to help construct theories or conceptual analyses.

It will be helpful to coin some rough and ready terminology that expresses these distinctions. Let's first differentiate “folk” from “philosophical” intuitions:

Folk Intuitions: the intuitions of theoretically naive people.

Philosophical Intuitions: the intuitions of theoretically sophisticated people.

It is important to note a difference in these kinds of intuitions. Folk and philosophical intuitions are domain specific. That is, in order to have philosophical intuitions one must have a sufficient level of theoretical training in that domain. This usually requires a certain amount of time spent reflecting on the issues. This reflection can take various forms, such as sorting out intuitions that one might have spontaneously but that one later finds untenable for theoretical reasons, or testing intuitions against theories and other intuitions to see which intuitions survive. In short, one does not get philosophical intuitions on the cheap—one has to work to get them.

It will also be helpful to make a general distinction between two types of philosophers.

Traditional Philosophers: Philosophers who think the only kind of intuitions that have any evidential weight for philosophical purposes are philosophical intuitions.

Experimental Philosophers: Philosophers who think both philosophical and folk intuitions have evidential weight for philosophical purposes. Both types of intuitions reveal essential features of concepts and implications of theories that are not revealed by either type alone.

The important distinction between traditionalists and experimentalists is that the former think that *only* philosophical intuitions are relevant to philosophy. For the traditionalist, folk intuitions are only relevant in a supporting role. For example, philosophical intuitions may gain additional support from the fact that most folk also think that way. However, to the extent that folk and philosophical intuitions diverge, traditionalists believe that folk intuitions can be ignored. Experimentalists contend that it is unlikely that extended reflection on philosophical intuitions will reveal all essential components of some central philosophical concepts. According to the experimentalist, folk intuitions are necessary parts of philosophical practice.

Interestingly enough, there is a seeming distinction in the way experimentalists use folk intuitions. Experimentalists can be further divided into two groups.

Critical Experimentalists: Experimentalists who think: (1) Folk intuitions reveal that

philosophical intuitions are inaccurate, (2) philosophical intuitions do not have a special status over and above folk intuitions, and the shortcomings of one express similar shortcomings of the other, or (3) philosophers wrongly identify what intuitions the folk have and thereby cannot use them as evidence for their theories or conceptual analyses.

Constructive Experimentalists: Experimentalists who think folk intuitions give us insight into philosophical concepts and theories that cannot be gained by philosophical intuitions alone. A fully formed conceptual analysis or philosophical theory depends at least in part on folk intuitions.²

It should be noted that both critical and constructive experimentalists are critical of philosophical intuitions, but in different ways. They both argue that philosophers need empirical evidence that the folk have the intuitions that philosophers ascribe to them. But, some critical experimentalists attempt to use folk intuitions as a proxy to show that philosophical intuitions are not good sources of evidence. Constructive experimentalists think that folk intuitions give us unique access into concepts and theories that philosophical intuitions alone cannot provide, even if philosophical intuitions can still be valuable sources of evidence.

My guiding question for this dissertation is the following one: Why should philosophers care about folk intuitions? After all, nobody denies that folk intuitions are often mistaken. In fact, there is very good experimental evidence that most folk sometimes make systematic errors (Gilovich & Griffin, 2002). In certain circumstances, some intuitions that the folk have are systematically wrong. Why, then, should we think that folk intuitions are philosophically important if we know they are sometimes systematically wrong? While we are not sure that all or even most folk intuitions are systematically wrong, they are wrong in some well defined situations. If they are systematically wrong in some situations, then there is at least some concern that they will not be good sources of evidence for philosophical theories and conceptual analyses.

We cannot be sure if folk intuitions about philosophically important issues are correct. We cannot be sure because we do not know what the right answer is to many philosophical questions. Therefore, it is not clear how to determine whether folk intuitions about philosophically important issues are systematically wrong. However, both the constructivist and the critic think that there is a property of folk intuitions that makes them relevant to philosophical debates—namely, the reliability of folk intuitions. The notion of reliability used here is modest

2 These distinctions roughly follow some distinctions made by Nadelhoffer & Nahmias (2007).

and need not have any connection with truth. Rather, reliability here is more akin to a notion of stability—intuitions are reliable if they are not influenced by extraneous factors. For example, take a scenario *S*. An intuition is reliable if, under normal conditions, one's intuition about *S* does not change in response to extraneous factors. To illustrate, an intuition about *S* should not change under normal conditions depending on when *S* is presented. Because the time when *S* is presented is an extraneous factor, if one's intuition does change, then we can say that intuitions about *S* are unreliable.

If all folk intuitions about philosophically important issues are systematically unreliable, then at least part of the critic's thesis holds—folk intuitions are unreliable and may indicate a similar unreliability in philosophical intuitions. However, if some folk intuitions about philosophically important issues are reliable, then one cannot make the inference to the unreliability of philosophical intuitions. Likewise, constructivists should be aware of the reliability of folk intuitions. If some folk intuitions are unreliable, then they cannot be used for theory formation or conceptual analyses. Folk intuitions cannot be used because if some intuitions are unreliable, then it is not clear which intuitions are to be used for theory formation or testing conceptual analyses. For example, at time *T*₁, John has the intuition that Smith knows in the Gettier style case. Then, because of some extraneous factor, at *T*₂ John has the intuition that Smith does not know. Because John's intuition is influenced by an extraneous factor and because we do not have a definitive answer to whether Smith knows, we have no non-arbitrary way to pick which of John's intuitions are to count as evidence.

By the very nature of the experimentalists' approach, it seems as if they must shoulder additional burdens that traditionalists do not. Experimentalists, because they rely on empirical data, can only legitimately make inferences inside the domain in which they collect their data. For example, experimentalists may find that folk epistemic intuitions are unreliable, but it would be improper to generalize that unreliability to all intuitions or to intuitions in another domain. Hence, the reliability of intuitions in one domain says nothing, empirically, for or against the reliability of intuitions in another domain. However, traditionalists defend philosophical intuitions in general. That is, traditionalists maintain that philosophical intuitions are legitimate sources of evidence across domains because they are the result of prolonged reflection. The reflective feature of intuitions is what traditionalists claim gives philosophical intuitions the appropriate character to be used as evidence generally. Hence, the argumentative burden is on the “newcomers” to show the relevance of folk intuitions to philosophy in general.

In Chapter 2, I take a look at some arguments that attempt to establish, a priori, that philosophical intuitions are in fact reliable (Bealer, 1998, 1996, 1992; Reid, 2002, 1997; Sosa, 1998; Tidman, 1996; Williamson, in press). I argue that these arguments either do not show that philosophical intuitions are reliable or that such reliability entails that humans may be incapable of having philosophical intuitions. I do not argue, however, that humans do not have reliable philosophical intuitions. Rather, I only argue that it cannot be established a priori that human philosophical intuitions are reliable. If my argument is right, then two empirical problems arise. First, some empirical work must be done to help us understand in which fields intuitions are reliable. Second, given that the reliability of intuitions may be domain specific, we would like to know in what domains humans possess the capacity to have reliable intuitions. In an attempt to shed light on these two problems, I will explore four fields: epistemology, ethics, free will, and action theory.

In Chapter 3, I briefly review the role of intuitions in some theories of knowledge (Bealer, 1996; Goldman, 2000; Goodman, 1955). I present some arguments based on folk epistemic intuitions designed to highlight that there are some worries associated with supporting epistemological views with philosophical intuitions (Stich, 1998, 1990; Weinberg, Nichols, & Stich, 2001). I present data that support the critic's claim that folk intuitions are not reliable. However, I also report a surprising stability of folk epistemic intuitions. It is true that “the” folk do not have reliable intuitions because it does not appear that there is any “the” folk. Rather, there are groups of people who express stable epistemic intuitions. This stability is argued to be sufficient for constructivist uses of epistemic intuitions.

In Chapter 4, I look at two ways in which intuitions are used in ethics. The first is an attempt of philosophers to ground conceptual analysis in everyday, common-sense intuitions (Smith, 1994; Shafer-Landau, 2003). These everyday intuitions are claimed to indicate platitudes about ethical concepts. It is claimed that any adequate analysis of ethical concepts must do justice to the platitudes about moral concepts. I review some objections based on empirical data. These objections state that philosophers who make use of platitudes to guide their theories have not done sufficient empirical legwork to see what moral intuitions the folk have (Nichols, 2004a). What is required of those who make use of platitudes is some evidence that the platitudes these philosophers identify really are those indicated by intuitions of most folk. The second use tests ethical theories and analyses of moral concepts with intuitions. Critics claim that moral intuitions are not reliable enough to perform that service. However, I present data that

suggest that moral intuitions are in fact reliable. Again, while there is no “the folk,” there are groups of people who have reliable moral intuitions. This reliability is argued to be enough for folk intuitions to be used to test ethical theories and analyses of moral concepts.

In Chapter 5, I look at some issues surrounding folk intuitions and the free will debate. In the free will debate, some philosophers attempt to make appeals to folk intuitions in order to support their theories. These philosophers argue that if the folk have intuitions that favor their theory over another (e.g., compatibilism over incompatibilism), then that gives additional weight to their arguments (Campbell, 1951; Ekstrom, 2002; Kane, 1999; O'Connor, 2000; Pereboom, 2001; Pink, 2004; Smilansky, 2003). The problem, experimentalists point out, is that most philosophers who appeal to folk intuitions do not provide the requisite evidence that folk intuitions in fact support their view (Nahmias, Morris, Nadelhoffer, & Turner, 2004, 2005, 2006; Nichols & Knobe, in press). I review some of the literature that provides empirical evidence about the unreliability of folk intuitions. However, I again argue that there is stability in folk intuitions. Again, there is no folk wide stability in intuitions, but there is stability in groups of folk. This stability is sufficient to use folk intuitions in ways the constructivists want.

Chapter 6 reviews intuition's role in action theory. Experimentalists think that folk intuitions about intentional action are important for philosophical accounts of intentional actions for at least two reasons. First, some traditionalists think their alleged analysis of intentional action squares with folk intuitions about intentional action (McCann, 1998). However, the critics contend, what intuitions the folk actually have is an empirical matter that requires empirical evidence. The critic points out that no such evidence is provided by many traditionalists (Nadelhoffer, 2006a). Second, constructivists think that folk accounts of intentional action are important for philosophical accounts of intentional actions, and the folk accounts reveal features of intentional actions that are left out of some philosophical treatments of intentional action (Knobe, 2003a, 2003b, 2004a; Knobe & Burra, 2006; Nadelhoffer, 2004a, 2004b, 2004c). For example, the moral status of a behavior can influence whether one is willing to say the behavior is intentional. These results, experimentalists suggest, are significant because in some sense we want our philosophical analysis to be able to cover these phenomena. After all, if moral factors influence folk intuitions, then we should want to know either why we can ignore the phenomenon (Adams & Steadman, 2004a, 2004b) or we should develop an analysis that can accommodate the phenomenon. The same question from previous chapters is repeated here—how reliable are folk intuitions, and are they reliable enough so that traditional philosophers

should pay attention to them? Again, reliability of intuitions is found at the group level.

Chapter 7 draws conclusions about the uses of folk intuitions. Given that intuitions in all four domains are found to be reliable, we should have some degree of confidence that folk intuitions can be used to test some philosophical theories and conceptual analyses. Specifically, I argue for the Fragmentation of Folk Intuitions thesis that states if folk intuitions are to be used as evidence, then some folk intuitions must be reliable. Because there is evidence that intuitions reliably indicate something, the Fragmentation of Folk Intuitions thesis is true. I conclude that the Fragmentation of Folk Intuitions thesis is sufficient for constructivist purposes.

CHAPTER 2: A PRIORI DEFENSES OF INTUITIONS

The dominant philosophical methodology in 20th century analytic philosophy attempts to offer analyses of various concepts. These analyses are then tested from the armchair. Philosophers test these analyses by coming up with counterexamples to either the necessity or sufficiency of the conditions. The success of these counterexamples depends on whether one intuitively thinks that the cases falsify the necessity or sufficiency of the conditions offered. For example, most people think that Gettier cases satisfy the three jointly sufficient conditions of a justified true belief account of knowledge. However, most people also have the intuition that the person in the Gettier cases does not have knowledge. Hence, the intuitions these philosophers have show that the conditions offered by the justified true belief account of knowledge are not sufficient for knowledge. This analysis-counterexample methodology is typical of contemporary Western analytic philosophy.

In recent years, the use of intuitions in philosophy as measures for adequate theories or analyses of concepts has been called into question by both philosophers and psychologists. Psychologists have performed studies that suggest humans are not very good intuitive reasoners, and intuitions may systematically vary between groups of people. These studies indicate that intuitions may not be reliable. For example, studies show that many intuitive judgments are wrong and that intuitions can be manipulated. For example, some studies suggest that people do not reason in accordance with the rules of probability (Tversky & Kahneman, 1982). Likewise, there are studies that suggest people's intuitive judgments are sometimes illegitimately influenced by factors of which they are not aware (Nisbett & Wilson, 1977). These studies give some reason to think that philosophers also fall victim to similar errors or manipulations. If it is true that no intuitions are reliable, philosophers' intuitions included, then it seems as if the common methodology in the philosophical tradition is in trouble. After all, if intuitions are as variable and unreliable as these studies suggest, then it is not clear which intuitions are supposed to be evidence for theories and conceptual analyses.

Of course, there are those who defend the use of intuitions in philosophy. One defender, George Bealer (1992, 1996, 1998), argues that intuitions are not only reliable indicators of truth,

they are necessary for the philosophical endeavor. He argues that when intuitions go through the correct procedure—reflective equilibrium—then those intuitions are reliable sources of evidence. Paul Tidman (1996) argues that intuitions are “innocent until proven guilty.” Because there is no non-self defeating way to show that intuitions are not good sources of evidence, they remain legitimate sources of evidence. Ernest Sosa (1998) gives a modified version the “innocent until proven guilty” strategy. His main contention is that those who use the empirical evidence that we are poor intuitive reasoners overreach. In addition, Sosa argues that the type of evidence intuitions provide is identical to the types of evidence that other belief forming mechanisms such as perception and introspection provide. Because we accept the latter as sources of evidence, we should also accept the former.

In what follows, I review arguments offered by Bealer, Tidman, and Sosa that intuitions are good sources of evidence for philosophical theories. I argue Bealer's, Tidman's, and Sosa's arguments are not sufficient to establish a priori that intuitions are good sources of evidence for philosophical theories. First, Bealer's reliance on reflective equilibrium as a means for discriminating good intuitions from bad intuitions may be flawed. Second, Bealer's view that intuitions *can* serve an evidentiary base for justifying beliefs is too weak to guide some of our actual philosophical endeavors. Third, I argue that Sosa's and Tidman's arguments do not establish a strong enough conclusion that, a priori, intuitions are reliable sources of evidence for philosophical theory. Finally, I distance myself from radical skepticism about intuitions and deflate the “self-defeating” argument that some have posed to those critical of intuitions (Bonjour, 1998; Kaplan, 1994; Siegel, 1984; Williamson, in press).

1. Bealer's Approach

According to Bealer, philosophical intuitions are special entities. On his view, philosophical intuitions meet the following conditions: (a) they are the result of conscious effort, (b) the process by which one comes to have philosophical intuitions is at least minimally under one's control via effort and attention, (c) a multi-stepped process may be involved in reaching the final philosophical intuition, and (d) intuitions are a priori, self-evident, basic knowledge (Bealer, 1996, p. 123-4). Philosophical intuitions are a distant cousin of the more common “folk” intuitions. Philosophical intuitions differ from folk intuitions because they result from critical reflection while folk intuitions are merely “uncritical beliefs” (Bealer, 1998, p. 202).

Given this characterization of philosophical intuitions, Bealer argues that intuitions are valid sources of evidence. Of course, as we are all painfully aware, not all intuitions track the

truth. For example, Bill may have the intuition that Smith knows in a Gettier style case and Jane may have the intuition that Smith does not know. Obviously, they can't both be right. If people have conflicting intuitions, then the “good” intuitions must be sorted from the “bad” intuitions. Bealer calls the process which sorts and uses intuitions as evidence the “Standard Justificatory Procedure” (SJP). The SJP consists of the following process:

- (1) Canvassing intuitions; (2) subjecting those intuitions to dialectical critique; (3) constructing theories that systematize the surviving intuitions; (4) testing those theories against further intuitions; (5) repeating the process until equilibrium is approached. (Bealer, 1996, p. 122)

Intuitions that survive the SJP gain the status of philosophical intuitions. They are no longer uncritical intuitions of the masses but rather are intuitions that have survived philosophers' critical reflection and attention. Thus, they acquire an epistemic status that folk intuitions lack.

On Bealer's view, the special epistemic status of philosophical intuitions allows them to be used as evidence for philosophical claims. This is so because philosophical intuitions have a special modal tie to the truth. Bealer calls the view that accounts for this tie modal reliabilism—a view he characterizes in the following way:

For suitably good cognitive conditions, it is necessary that, if while in such conditions a subject goes through the whole procedure of *a priori* justification [SJP], then most of the propositions derivable from the resulting comprehensive theoretical systematization of the subject's intuitions would have to be true.³ (Bealer, 1996, p. 130)

“Suitably good cognitive conditions” are conditions that are ideal for theory formation. When these conditions are met, one is suitably positioned to systematize one's intuitions. For example, conditions that generate cognitive illusions or performance errors would be absent in good cognitive conditions. Because these intuitions have a strong tie to the truth, one should be able to solve most of the questions that have plagued philosophy. After all, given the appropriate cognitive conditions and the SJP, most of one's theoretically informed philosophical intuitions, unlike pretheoretical folk intuitions, could not lead to false judgments.

1.1 The Conceptual and Empirical Case against Intuitions and the SJP

In an interesting research program, Richard Nisbett (Nisbett et. al., 2001) has discovered

³ Bealer considers a weaker version of modal reliabilism. The weaker version states, “...if someone in those cognitive conditions were to process theoretically the deliverances of the candidate source, the resulting theory would provide a correct assessment as to the truth or falsity of most of those deliverances” (Bealer, 1998, p. 219). I don't expect that this weakening will affect the main thrust of my argument.

that people of different cultures reason differently about the world. Stephen Stich argues that this kind of cognitive diversity is a problem for the proponent of intuitions and the SJP. Cognitive diversity describes the plurality of ways people form and update their beliefs (Stich, 1998, p. 96). One way cognitive diversity may occur is by systematic variations of people's intuitions about epistemic justification. It might turn out that some people intuitively accept beliefs as justified that are the results of known fallacies. For example, some people wrongly think that the probability of a conjunction is greater than one of the conjuncts (Tversky & Kahneman, 1982). If this is right, then people can accept principles that are wrong; yet, they are seemingly in reflective equilibrium for them. But just because principles are in reflective equilibrium does not mean that one is justified in believing them (Stich, 1998, p. 100). That would entail that one can be justified in believing a fallacy, which many would hesitate to accept.

Cognitive diversity might also occur when there is more than one philosophically relevant concept in single domain. If there is conceptual diversity, then relativism might come about because what fits one group's concept may not fit a different group's concept (Stich, 1990, p. 88). For example, there may be more than one concept of justification. Insofar as the theoretical systematization of intuitions fits well with the concept, that systematization is thought to be true by those who have the concept. If that is true, then the SJP does not help us determine which theoretical systematization is the right one because there may be a plurality of “right” ones. Concepts of justification may be idiosyncratic consequences of particular environments and upbringings; hence, it seems as if the defender of intuitions and the SJP is in hot water if she cannot rule out these possibilities (Stich, 1990, p. 95).

Stich's arguments are fairly easy to take in stride because they rest on the *possibility* of cognitive diversity and different concepts. Building off Stich's earlier work, Jonathan Weinberg, Shaun Nichols, and Stich (WNS) (2001) argue that the use of reflective equilibrium, with intuitions serving a critical role in the process, might generate relativism in the actual world. To establish the systematic variation of intuitions, WNS provide scenarios to the different groups of “folk” that describe a person in various epistemic situations. One scenario they use is a typical Gettier case. Responding to the Gettier style case, about 25% of Westerners, 55% of East Asians, and about 60% of Indians thought the person really knows (Weinberg et al., 2001, p. 443-4). These studies indicate that there is systematic diversity of epistemic intuitions about Gettier style cases. Hence, there is a real risk that theories that depend on intuitions as input really will give different normative results (Weinberg et al., 2001, p. 445). At a minimum, WNS think that the

burden is shifted to those who defend intuitions and the SJP to explain why these empirical results are not threatening. Specifically, if the intuitions used as inputs into the SJP can lead to different theories, then one must either (a) allow that philosophical theories founded on intuitions and the SJP could be hopelessly relativistic, (b) prove that intuition based theories do not lead to relativism, or (c) concede that intuitions cannot be used as evidence (Weinberg et al., 2001, p. 447-8).

1.2 Empirical Pressure on Bealer's View

Bealer could respond to these worries in at least two ways. First, the possibility of people arriving at different theories given different intuitions as input is not threatening to his view. We have good evidence that many of our intuitions overlap on many cases; thus, the mere possibility of variation should not be enough to scare us. Second, as a reply to WNS, Bealer could simply say that the type of intuitions in which he is interested—philosophical intuitions—is simply of a different kind than the intuitions with which WNS are concerned. WNS are concerned with folk intuitions, and Bealer is concerned with philosophical intuitions. Hence, neither set of intuitions have much to say about the other.

Bealer seemingly has replies to both challenges. Why, then, isn't that the end of the story? The reason is that empirical evidence indicates humans are not the kinds of creatures who have the requisite philosophical intuitions to engage in philosophy as Bealer conceives it (see section 1.3). If *we* do not have the right kinds of intuitions, then *we* are not the kinds of beings who can solve philosophical questions in the way Bealer thinks we can.

1.3 The Problem of Relativism

The results presented by WNS are significant because they suggest at least two things—namely, that there can be more than one seemingly coherent set of beliefs and that these different sets of beliefs can be in reflective equilibrium. The question is which system is to be preferred if the beliefs are in reflective equilibrium with the intuitions used to generate them? Prima facie, I think that Bealer is in a dilemma. On the one hand, it seems that Bealer cannot say any set of intuitions that passes the SJP has a strong modal tie to the truth because it is possible that two competing systems that pass the SJP can be in reflective equilibrium. If the two systems are competing, then it cannot be the case that most of the pronouncements of both systems are necessarily true. On the other hand, there might be something in addition to the SJP that gives the strong tie to the truth. If that is the case, then Bealer must explain what the additional element is.

One element in addition to the SJP may be that the input data are philosophers' intuitions and not merely error prone folk intuitions. But even if one looks at philosophers, one finds diverging intuitions. There are several *prima facie* examples of competing intuitions among philosophers. Consider, for instance, the conflict between deontologists and consequentialists.⁴ They often have different intuitions about the moral status of a particular action. For example, a consequentialist may think it is permissible to falsely accuse an innocent man to prevent a riot, but a deontologist might think that action is impermissible. After reflection, the consequentialist and deontologist make different judgments about the moral status of some actions.

If both deontologists and consequentialists use the SJP to generate their theories, then in some cases deontologists and consequentialists have different philosophical intuitions concerning what the morally correct action is. Otherwise, it would be hard to explain the divergent prescriptions. People who espouse either consequentialism or deontology have spent a great deal of time thinking about the issues, and their intuitions should count as philosophical if any do. If these two views are based on philosophical intuitions which have passed through the SJP, then it seems fairly unlikely that convergence between these two views is going to be achieved. And, because the two views give conflicting prescriptions, it cannot be that they both have a strong tie to the truth. At least one of the views must be wrong on Bealer's view. However, it is hard to see how to decide between them if both views are in reflective equilibrium with the philosophical intuitions that generate them. Once again one is faced with the problem of relativism much like the one suggested by WNS.

One might think that the relativism entailed by Bealer's view is not pernicious. Maybe some groups of people have intuitions that are in reflective equilibrium and yield incorrect results, and there are other groups of people whose intuitions are in reflective equilibrium and yield correct results. People's intuitions may be in reflective equilibrium and yield the incorrect results because they may not have a wide enough range of intuitions, or the theories with which they attempt to achieve equilibrium are false. Thus, we should only pay attention to people whose intuitions are linked to the correct results.

The relativism entailed by Bealer's view is not pernicious only if one of the following two possibilities is true. First, the relativism only obtains in situations where we can independently determine the correct answer. If we know the correct answer, then we can check and see if intuitions match those standards. However, in most philosophical domains we are not sure what

4 I thank Kim Sterelny for bringing this particular example to my attention.

the correct answer is. One is a deontologist or consequentialist *because* one finds those views to be intuitively correct after prolonged reflection and not because we independently know what the right answer is. If these theories are the result of the systematization of philosophical intuitions, there is no basis to prefer one set of intuitions over another. We just are not sure which view is correct, so we cannot tell which set of intuitions is correct.

Second, one might think that the relativity of intuitions is not problematic because we should appeal to the intuitions of experts rather than the intuitions of laypeople. However, as Stich argues, if one does not assume that experts' intuitions are in fact correct (an assumption which is obviously question begging), then it seems possible that experts' intuitions are every bit as off as the *hoi polloi* (1993, p. 86). Thus, appealing to expert authority does not help matters much—especially when the experts themselves often have diverging intuitions!

Perhaps when there are divergent intuitions, the concepts are not adequately grasped. According to Bealer, two people who satisfy the appropriate cognitive conditions will understand a concept in the same determinate manner. In turn, both will have the same intuitions that are related to that concept. Bealer thinks that if one understands a concept determinately, then almost all intuitions that have to do with that concept are true (1998, p. 203). A concept is possessed determinately when the following two conditions are met:

1. The subjects possess some propositional attitude toward the proposition that has the concept as its conceptual content.
2. The subjects do not misunderstand the concept or have incomplete knowledge of the concept (Bealer, 1998, p. 221-2).

When one possesses a concept determinately one *really* knows the concept fully and completely. Hence, if two people possess the same concept determinately in ideal conditions, then differences of intuitions simply are not possible.

Determinate concept possession could play a role in answering the relativity problem in at least three ways, all of which are problematic. The first way is if people already have determinately held concepts prior to engaging in the SJP. One way people could already determinately possess a concept is that the concept makes up part of some of our basic competencies. For example, our basic competencies with '2', '+', '4', and '=' ensure that we all have the intuition that $2+2=4$. Likewise, we may determinately possess some philosophically interesting concepts that guarantee most of our intuitions involving those concepts are true.

The problem with the first way of answering the relativity problem arises in two stages.

First, this way makes the SJP impotent. It's not that intuitions are somehow imbued with a strong tie to the truth because they pass the SJP, but they pass the SJP because they are already tied to the truth. To see the problem, assume that all intuitions before going through the SJP are folk intuitions, and all surviving intuitions are philosophical intuitions. Let's say that Jane has a folk intuition that p . Call this intuition (F). Let's also say that Jane possesses concept (C) determinately, and F applies to C . That means that F already has a strong tie to the truth given that Jane holds C determinately. Now imagine that Jane uses the SJP on F . What possible difference could the SJP make on the connection of F with the truth? It seems none. Certainly, using the SJP can help us *access* determinately possessed concepts by eliminating interference, simple mistakes, and inattention. But the SJP does not have any bearing on the *truth* of F . In this situation, the only property that F gains is that it went through the SJP where that property has no effect on the truth of F .

Second, if the only property that F gains is that it passes through the SJP and that property has no bearing on the truth F , then intuitions are true simply because of their relation to a determinately possessed concept and nothing else. Given suitably good cognitive conditions and a determinately held concept, one's intuitions already have a strong modal tie to the truth. The SJP does nothing to strengthen or ensure that tie. If this is right, then determinate concept possession simply solves the problem of conflicting intuitions by begging the question. The question is whether humans possess enough philosophically relevant concepts determinately. To claim that possessing a concept determinately entails from the very outset that most intuitions involving that concept are true begs the question against those who think that humans do not possess enough philosophically relevant concepts determinately.

The second way is that intuitions are weeded out by the SJP that uses determinately possessed concepts. Some intuitions are bad and others are good, and the SJP using determinately held concepts is what sorts the good intuitions from the bad ones. Given that the SJP is a process by which one balances one's intuitions with, among other things, determinately possessed concepts, determinately possessing a concept will rule out all intuitions that do not have a strong tie to the truth.

If Bealer makes this move, then he simply begs the question once again. What is required in addition to possessing a concept determinately is that there is only *one* relevant concept in that domain. If there is more than one concept in that domain, then people might possess different concepts determinately and thereby have different intuitions surviving the SJP. For example,

there may be more than one concept of epistemic justification or moral rightness. If there is more than one concept of justification or moral rightness, then people may possess those different concepts determinately and yet have different intuitions about what constitutes right action or epistemic justification. Recall the example of a deontologist and consequentialist. Presumably, a consequentialist and a deontologist possess most of the relevant ethical concepts determinately; yet, they still have differing philosophical intuitions about what constitutes some right actions. This might indicate that consequentialists and deontologists employ different concepts of rightness that influence their judgments about right actions. It is unclear why simply possessing a concept determinately would solve the problem of conflicting intuitions if there is more than one concept in the domain.

The third way is that the SJP helps along the way to determinately possessing a concept. That is, the intuitions that pass the SJP are the ones that point to determinately held concepts. If that is the case, then it seems that we have the same problem that Stich and WNS suggest. If we start with different intuitions and then attempt to systematize those until we think we have discovered the correct concept, the variation of input intuitions may affect the output concept. Not only is this a live possibility as Stich argues, but the empirical evidence from WNS adds additional pressure because there are systematic variations in actual people's intuitions.

Thus, Bealer's approach is subject to the worries presented by Stich and WNS. We cannot simply assume from the outset that we possess enough concepts determinately because that begs the question against those who think that we do not. Even with conjunction of the SJP and determinate concept possession there are problems. We cannot assume that there is only one relevant concept in a domain. If there is more than one concept, then there can be different, competing sets of intuitions that are thought to be true. Finally, we cannot assume that the SJP helps along the way to determinate concept possession because different input intuitions can result in seeing different concepts as being possessed determinately. Therefore, it is not the case that, a priori, the combination of SJP and determinate concept possession shows that humans are the kind of beings who can have the types of intuitions to support Bealer's view.

1.4 The Problem of Infinite Intelligence

One might think that when different philosophical intuitions lead to different theories we are not dealing with an ideally rational cognizer. What one should base philosophy on is the ideally rational cognizer under ideal conditions, where the intuitions that the idealized cognizer produces from the SJP are true. Bealer admits that the conditions under which one achieves an

ideal situation would be very high. A natural question for Bealer is how high must one set the bar? He has no definite answer to this question, but he does give us some idea. Bealer writes:

Most of the central questions of philosophy do not seem to be the sort of questions requiring infinitary intelligence (e.g., for doing infinitary proofs, infinitary computations, etc.); some finitary level (perhaps well beyond ours) ought to suffice...If this is right, the issue comes down to the question of what level of finitary intelligence would be required (for having a sufficiently wide range of intuitions) to yield Autonomy⁵...Intuitively, however, for any finite level of intelligence, it is possible for some being to be that intelligent. So, if there were a barrier to Autonomy, it would have to be something other than intelligence. (Bealer, 1996, p. 132)

There are two problems with Bealer's argument for philosophy's autonomy. First, how would we ever know when we have a sufficient level of finitary intelligence? Bealer must provide criteria for determining when one has achieved a sufficient level. If there is no way to tell when one is in sufficiently good cognitive conditions, then intelligence is a very real problem for the intuition approach. It may not be the case that philosophy requires infinite intelligence, but what if it requires a vastly greater intelligence than humans have? In both cases it seems that humans are simply ill-equipped to handle the questions that philosophy poses. To guarantee an intuition's tie to the truth, Bealer would have to concede that we might, to borrow an apt phrase from Stephen Stich, have to have a brain "the size of a bathtub," and that certainly would be too demanding (Stich, 1990, p. 153).

Bealer might reply that there is hope that philosophy can answer these problems. He claims that philosophy as a "civilization-wide project" might arrive at the same type of convergence that we find in science. In fact, science is not an individualistic project. Science needs many individuals over time in order to generate its theories, and some think there is gradual progress toward the truth. Even if there is not complete convergence on scientific theories, nobody is calling for an end of the current way science is practiced. Similarly, it might be premature to call an end to the intuitional approach to philosophy. What we need to do, Bealer argues, is simply give philosophy more time to work on its problems. There are many examples where intuitions in philosophy have converged, and that provides a basis for thinking that wide-scale convergence is also possible (Bealer, 1996, p. 139). For example, there is wide-scale

5 Bealer gives the following description of philosophy's autonomy: "Among the central questions of philosophy that can be answered by one standard theoretical means or another, most can in principle be answered by philosophical investigation and argument without relying substantially on the sciences" (Bealer, 1996, p. 121).

consensus that the conditions of the traditional Justified True Belief account of knowledge are insufficient for knowledge. This convergence gives us reason to think that, incrementally, we can slowly solve problems in philosophy.

I agree with Bealer that we “can” do philosophy carefully enough to answer “a substantial number of central philosophical questions.” But, that is a weak claim. It is *possible* for us to answer these questions, but what is needed in addition is evidence that it is likely that we will answer these questions. When we look at the history of the intuitional approach to philosophy, however, the prospects of convergence seem bleak. In order for Bealer’s claim to hold, almost all philosophical intuitions must converge on a single set of intuitions after a process of reflective equilibrium in order for philosophy to be autonomous. That seems to be a matter that can be answered empirically, and there is reason to think that Bealer’s a priori method will not be born out. First, WNS’s studies suggest there are large cross cultural and socio-economic differences in some philosophically relevant intuitions. Second, studies suggest philosophical intuitions are biased by non-evidential features. For example, in some cases intuitions about philosophically important issues are subject to order effects (Swain, Alexander, & Weinberg, 2008). Third, there is a great deal of evidence from other philosophically relevant domains suggesting a variety of biases, intuitions, and concepts that might entail intuition relativism.⁶ Therefore, if any of these are true, then Bealer’s a priori method is not one humans are well-suited to use to solve philosophical problems.

Second, Bealer claims that for any finite level of intelligence it is possible for finite beings to obtain it. But again, thinking that it is possible for humans one day to achieve such conditions is a very weak claim and not one that should ground our epistemic, or any other, endeavor. This possibility would seem to underwrite using philosophical intuitions to tackle any theoretical problem.⁷ For example, there are various examples in the history of philosophy where philosophers have used philosophical intuitions to solve problems of physics, and those solutions have been drastically wrong.⁸ It might be argued that Descartes used philosophical intuitions to

6 For a few examples, see Cushman & Mele, in press; Knobe, 2003a; Knobe and Burra, 2006; Mele & Cushman 2007; Nadelhoffer, 2006b; Nahmias, Morris, Nadelhoffer, & Turner, 2006; Nichols & Ulatowski, 2007.

7 In fact, Bealer suggests something like this in what he calls the Authority of Philosophy thesis, which states, “Insofar as science and philosophy purport to answer the same central philosophical questions, in most cases the support that science could in principle provide for those answers is not as strong as that which philosophy could in principle provide for its answers. So, should there be conflicts, the authority of philosophy in most cases can be greater in principle” (Bealer, 1996, p. 121).

8 One might worry that Descartes simply confuses a conceptual matter with an empirical matter. Intuitions, in

conclude that a vacuum does not exist. He thought long and hard about the vacuum and related concepts. Descartes had the intuition that extension is identical to space, body cannot exist unextended, and a vacuum is by definition space with nothing in it. Because of these intuitions, there simply cannot be a vacuum (Descartes, 1985, p. 230). The claim that space must be occupied by something is a modal claim that requires a philosophical intuition. But, as it turns out, Descartes was wrong and there can be space with nothing in it. Thus, mistaken intuitions result in a mistaken view about the vacuum.

If we are to follow the advice given by Bealer, then it seems that the possibility that one day we could find the correct answer about the vacuum warrants using philosophical intuitions to that end. The vacuum example suggests that the possibility that we could achieve the right level of intelligence to use only philosophical intuitions to solve many questions in physics is analogous to us achieving the right cognitive conditions to use only philosophical intuitions to answer many philosophical questions. But something over and above philosophical intuitions has helped us discover the nature of the vacuum. Likewise, it is reasonable that something over and above philosophical intuitions will be required for us to discover philosophically important things like the nature of knowledge. Hence, the mere possibility that we can find correct philosophical accounts using philosophical intuitions is not strong enough to warrant us relying on philosophical intuitions totally on a priori grounds.⁹ What we should want is some evidence that our intuitions have the kind of connection to the truth Bealer suggests they do.

Perhaps these concerns underwrite Bealer's concern that it might not be nomologically possible for humans to obtain such conditions. If it is not nomologically possible for humans to obtain the conditions necessary for our intuitions to be strongly tied to the truth, then there seems to be little reason for *us* to guide *our* epistemic endeavors according to Bealer's method. In order for the possibility to which Bealer appeals to guide our philosophical endeavors, it seems that we should require that there is some evidence that *we* can answer those questions. Given the history of philosophy, and the growing evidence of systematic diversity of intuitions, we have reason to think that although convergence is possible, it may not be likely for *us*.

Bealer's sense, are only about conceptual matters. So, the physics example is beside the point. However, Descartes engaged in physics in a purely conceptual way—his concepts of vacuum, space, and body are all discoverable in the same a priori way Bealer thinks philosophical concepts are discoverable. Second, it is hard to show a clear case where philosophical intuitions involving “purely” conceptual matters are *wrong*.

9 In fact, alternative approaches to epistemology add additional pressure to Bealer's approaches. For example, Michael Bishop and J.D. Trout's (2005) view does not necessarily rely on philosophical intuitions in a pivotal role.

Finally, one may wish to remain optimistic that the intuitional approach that Bealer characterizes is still the way to go. After all, Bealer's claims are modal claims. That is, it is possible that in some world there are creatures that achieve the appropriate cognitive conditions and have a wide variety of intuitions. It is possible that the resulting theories that these creatures produce do have a strong tie to the truth. Thus, one might think the problem of infinite intelligence and the problem of relativism are beside the point because the actual world does not provide a counterexample to the possibility that there can be creatures in such conditions. Even if *our* intuitions do not obtain strong ties to the truth, it is still possible that some creatures' intuitions do. As Bealer claims, one only needs the possibility that intuitions obtain strong ties to the truth in order to underwrite his view that intuitions can serve as evidence (1998, p. 203). Hence, we can continue working in the intuition tradition because there is a possible world in which all these favorable conditions obtain, and we should simply attempt to approximate that possible world as much as possible.

Certainly, nobody should deny it is possible that there are beings who are smart enough and are in the appropriate position to use intuitions as evidence. I don't find that claim particularly problematic. However, if we are to think the mere *possibility* of there being such entities is sufficient to underwrite using intuitions in philosophy, then we should also agree to the following point: It is possible that magic eight balls always give correct answers to philosophical questions.¹⁰ Indeed, depending on one's conception of modality, there is a possible world where the oracle of the eight ball gives its philosophical proclamations and all those proclamations have a strong tie to the truth. But to claim that *we* should use the pronouncements of magic eight balls to justify our philosophical theories is absurd. Likewise, the mere possibility that intuitions could give the correct answers to philosophical problems is insufficient for them to be used to justify our philosophical theories.

2. The Mint of Nature Argument

Another a priori defense of intuitions is the "Mint of Nature" argument (Reid, 1997, p. 168-9). Paul Tidman (1996) gives the following version of the Mint of Nature argument. We are justified in believing the results of any belief forming mechanism which we don't have reason to doubt (Tidman, 1996, p. 167). Initially, we are justified in believing the results of all belief forming processes because, at least initially, there is no evidence that the belief forming processes are defective. Moreover, initially there is no way to provide non-self-defeating

¹⁰ I owe this example to Michael Bishop.

evidence for any belief forming process (Tidman, 1996, p. 168). Tidman thinks this is so because in order to establish any track record for the belief forming process, one must use the process in order to develop a track record. Once a track record is established, one can evaluate the record to determine the process's reliability. For example, if one wants to show that one's visual perception is unreliable, one must use that mechanism to generate a track record. Then, one can make reference to other sense perceptions to show that the visual perceptions misfire. However, when it comes to one's intuitions the only way one can show that intuitions misfire is to use other intuitions. But that just is to assume that intuitions are reliable to show that they aren't—an obviously self-defeating argument. No general, non-self defeating arguments that intuitions are unreliable can be generated. Therefore, because all arguments against intuitions' reliability are self-defeating, intuitions are reliable (Tidman, 1996, p.169).

Of course, the Mint of Nature argument does not entail that intuitions can never be wrong. Rather, intuitions cannot be *systematically* wrong. As Tidman writes:

I am not suggesting that any particular intuitive judgment is immune from criticism. We can and do discover instances where our intuitions have misled us. We may come to learn that a basic faculty is in error by relying on the reports of our other faculties.

Moreover, it is conceivable that one could have reason to come to mistrust the deliverances of a particular faculty altogether. This would be the case if the outputs of the faculty were massively inconsistent with one another. If the deliverances of one faculty were massively inconsistent with the deliverances of another faculty, one would have good reason to reject one of the offending faculties. (Tidman, 1996, p. 169)

But, because we don't have reasons to doubt intuitions in general, we don't have a reason to doubt the deliverances of intuitions. Hence, until we have reason to think that intuitions are not reliable, we can continue to trust the deliverances of intuitions.

While the Mint of Nature argument may be interesting, it is not particularly compelling. The Mint of Nature Argument may justify believing some obvious truths. Tidman focuses on modal truths, such as the commonly used example that nothing can be red and green all over at the same time. When one reflects on the statement that nothing can be red and green all over at the same time, one has the intuition that the statement is true (Tidman, 1996, p. 161).¹¹ But some

¹¹ One might worry that we come to know that nothing is red and green all over at the same time on the basis of a sound argument. This might be true for some people. However, others (e.g., Bonjour) have claimed that they immediately see that the statement must be true. Hence, at least some people have the intuition that nothing can be red and green all over at the same time. Moreover, if we take Lewis's view of what an intuition is, then even a belief that is come to by reflection can count as an intuition. In any event, it looks like one can form the

care must be taken in generalizing from intuitions giving the right results in some cases to intuitions giving the correct results in all cases—a move which is obviously fallacious. The Mint of Nature argument does not offer a general justification of the use of intuitions in philosophy. Rather, the Mint of Nature argument gives us a *criterion* for determining whether intuitions are reliable. After all, the Mint of Nature argument tells us when using intuitions is permissible—namely, when we have no reason to doubt them. The question, then, is whether we have reason to doubt the intuitions.

The crucial premise, then, is that we don't have any reason to doubt the reliability of intuitions in philosophy. This premise might be false. First, WNS cast some doubt on the reliability of intuitions based on the systematic diversity of epistemic intuitions. If WNS are right, there is a worry that intuitions, even when they are systematized, are not reliable. If our intuitions are the result of idiosyncratic, local conditions, then that would seem to count against the idea that those intuitions are reliable. Hence, there is at least a worry that intuitions are not reliable.

Second, there is a well established psychological literature that shows that some quickly formed judgments are not reliable.¹² For example, there are a variety of framing effects that influence people's judgments. Framing effects present participants with logically equivalent scenarios; yet, given some manipulation, they illicit different judgments. There is a wealth of experimental data demonstrating framing effects. An everyday example will give a flavor of how frames can alter judgments. An experiment performed by Irwin Levin and Gary Gaeth (1988) suggests that judgments about the quality of ground beef are influenced by how the beef is described. Beef that is described as “75% lean” got significantly higher ratings of quality than beef that is described as “25% fat” (Levin & Gaeth, 1988, p. 376). Because the descriptions are logically equivalent, judgments about the quality of the beef should not be different—but they are. Hence, judgments issued in these cases are not reliable because the judgments about the beef change in response to extraneous factors. Because there is evidence that quickly formed judgments in some domains are not reliable, there is some reason to doubt that all quickly

judgment that it is true that nothing can be red and green all over fairly quickly (even if by sound argument), so it counts as an intuition on my account.

¹² On my view, the quickly formed judgments in the beef example may not be intuitions because judgments about beef are not the sorts that have normally be used in philosophy in the past half century. While I think this is true, I hope that the example is illustrative of how, because judgments about the beef are quickly formed, that intuitions could likewise be unreliable.

formed judgments are reliable. If this is so, then the responsible thing to do is to check and see if quickly formed judgments in other domains are reliable. This requires empirical evidence from the relevant domains. For example, if we find that quickly formed judgments (i.e., intuitions) about free will are not reliable, then we have reason to reject arguments based on those intuitions. However, if we don't find evidence that intuitions are not reliable, then we can have some confidence in intuitions in those domains.¹³

Third, the Mint of Nature argument may not be intended to establish the general reliability of intuitions. Rather, it may only be meant to establish domain specific reliability. For example, The Mint of Nature argument may only be intended to establish the reliability of intuitions about modal concepts and not the reliability of non-modal concepts in epistemology. If the Mint of Nature argument is meant to establish *domain specific* reliability, then it poses no challenge to testing whether or not people have reliable intuitions. That is, if we take this interpretation of the Mint of Nature argument, then it is consistent with the view that there is no general a priori defense of the reliability of intuitions in philosophy. We might not have reason to doubt intuitions about modal concepts, but we may have reasons to doubt intuitions about non-modal concepts. Hence, interpreting the Mint of Nature argument this way reduces the scope so that it is not wide enough to shield the intuition approach to philosophy from empirical investigation.

3. Sosa's Minimal Intuition

3.1 Sosa's Arguments for Minimal Intuition

Ernest Sosa also offers a priori arguments for the general use of intuitions in philosophy. He offers the following analysis of intuition:

(I) At t , it is intuitive to S that p iff (a) if at t S were merely to understand fully enough the proposition that p (absent relevant perception, introspection, and reasoning), then S would believe that p ; (b) at t , S does understand the proposition that p ; and (c) the proposition that p is abstract.¹⁴ (Sosa 1998, 259)

According to Sosa, (I) has a number of advantages because it only states what one *would* believe

13 Of course, just because we don't find any evidence that intuitions are unreliable in some domains does not mean that they are reliable. However, if the Mint of Nature argument is right, if we find that intuitions in some domains are reliable, we would have good reason to think intuitions are reliable in other domains. At a minimum, not finding evidence that intuitions in that domain are unreliable gives us *some* reason for thinking that they are reliable in that domain.

14 Concerning abstract, Sosa writes "I do not define what it is for a proposition to be abstract. Fortunately, our working grasp of the concept seems good enough for the present purposes" (1998, p. 258).

given the appropriate circumstances. For example, (I) does not commit one to thinking there are necessary truths, it does not commit one to a kind of Platonism, and it does not rely on being acquainted with the object in question (Sosa, 1998, p. 260). Also, any plausible contemporary account of intuitions does not commit one to the infallibility of intuitions. After all, Sosa agrees that some of the empirical data do show that humans sometimes have intuitions that are wrong (1998, p. 261).

Sosa defends the use of intuitions in philosophy with two analogical arguments. The first I'll call the So What Argument. The So What Argument runs as follows. We can agree that intuitions are sometimes wrong even in conditions that are conducive to having the right kinds of intuitions. However, just because we sometimes have intuitions that are wrong, so what? That does not mean that intuitions in general are wrong. And the empirical evidence only shows that our intuitions are sometimes wrong under some conditions and not that they are always wrong. Sosa thinks that this would be like saying that because we sometimes have some false perceptual beliefs we should discount beliefs based on perception—a move that nobody is willing to make. Because we wouldn't discount all perceptual beliefs on the basis of sometimes having false perceptual beliefs, we also should not throw out all intuitions just because some intuitions are wrong. Hence, the empirical evidence against intuitions does nothing to cast doubt on the general use of intuitions (Sosa, 1998, p. 261).

Sosa's second argument defends (I) against the charge that it only results in trivialities such as nothing can be red and green all over, and that it cannot support using intuitions for justifying philosophical positions. That is, critics of (I) might require in addition to (I) a detailed account of the mechanism that generates these intuitions and some reason to think that the mechanism is reliable (Sosa, 1998, p. 262). Sosa thinks that requiring such an account is too stringent. He asks us to suppose that we know some abstract proposition, like $2+2=4$, and we do so non-accidentally. Let's also suppose that one non-accidentally knows the visual proposition that there is a baseball flying at one's head. In this case, in order that the belief that a baseball is flying at one's head to count as knowledge the belief must be non-accidentally true. So, Sosa thinks that both perceptual and intuitional beliefs can be given the following general schema:

S Φ 's that p only if S believes that p in virtue partly of these facts: (a) that S understands the proposition that p, and (b) that the proposition that p is true and of a certain sort s, one appropriate for Φ -ing. (Sosa, 1998, p. 263)

According to the schema, there are elements in addition to just understanding p that are required

in order for one to appropriately Φ . Because we can substitute intuition and perception into Φ , there are contingent conditions in place on both introspection and intuition in addition to clearly understanding the proposition p (Sosa, 1998, p. 264).

Because there are contingent conditions in place on both introspection and intuition, and in some conditions one can both have perceptual and intuitive failures, “one must be gifted with a sensitivity to propositions in that field whose truth one mirrors or tracks” (Sosa 1998, 265). That is, according to Sosa, “with regard to intuition and with regard to introspection, we seem able to tell the limits of our abilities. Somehow we can tell the sorts of propositions that lie within the proper scope of our respective faculty, and to believe accordingly” (1998, p. 265). Because we can do this with respect to perceptual knowledge, and perceptual and intuitional knowledge fall under the same schema, by “parity of reasoning” if we are to accept the dictates of perception when they are restricted to the appropriate boundaries we must also accept the dictates of intuitions when they are restricted to the appropriate boundaries (Sosa 1998, p. 267). We don't need to know how we create these boundaries. We only need to make these boundaries sometimes. We have a good sense of where our intuitions are likely to be wrong and we have a pretty good sense of where our intuitions are likely to be right, and we make the boundaries accordingly. Because we do seem to make these boundaries sometimes, we don't need a detailed account of the mechanisms that generate intuitions in order for intuitions to be sources of evidence in philosophy.

3.2 Some Problems with Sosa's Minimal Intuition

There are a number of replies to Sosa's defense of minimal intuition. First, it seems that Sosa underestimates the force of the empirical arguments against intuitions. He claims that while the empirical evidence shows that sometimes intuitions are wrong, that alone does not mean that intuitions in general are wrong. That certainly is true. After all, it seems plausible that our intuition that $2+2=4$ is right. However, given that we know that intuitions are wrong in some areas, it casts *some* doubt on the use of intuitions in other domains. For example, let's think of the problem the other way around. Let's assume that the intuition that $2+2=4$ is right. Is that enough to underwrite the claim that intuitions in general are reliable? Of course not. As Sosa freely admits, intuitions in some other domains can be wrong.

So, where does that leave us? It looks like it leaves us no further along than before. After all, $2+2=4$ is not a philosophically interesting proposition, and is not one that is part of a live (well, at least not for most) philosophical debate. However, there are other domains where there

are hotly contested issues—action theory, free will, epistemology, ethics—where it is not clear that our intuitions tell us anything reliably. In these domains it seems perfectly acceptable to undertake an empirical investigation into intuitions as psychology has done with human reasoning. Hence, while the empirical arguments are not sufficient to warrant rejecting intuitions in all philosophically relevant domains, so too is Sosa's argument insufficient to warrant accepting intuitions in all philosophically relevant domains. What is required is some empirical evidence that intuitions in these domains are of the right sort Sosa specifies.

As for Sosa's second argument, there again is something certainly true about what he says. It seems true that one need not have a well-worked out account of the mechanisms that generate intuitions in order to say that we are justified in believing some things intuitions tell us. After all, we only have to restrict the boundaries somehow and only believe the pronouncements of intuitions inside that range. But again, we can cast some doubt on the mechanisms that generate intuitions and our ability to identify the range over which intuitions are reliable. And again, just because intuitions are reliable in some domains does not mean that they will be reliable in all domains. What is required is some evidence that allows us to decide one way or the other—after all, in some philosophically interesting domains it is not clear if any intuitions are reliable.

Moreover, it is not clear that people are in general able to create boundaries around where intuitions are reliable—especially in philosophically relevant domains. For example, research by Richard Nisbett and Timothy Wilson (1977) suggests that often we are not very good at understanding what exactly influences our intuitions about particular cases. That is, many factors that influence intuitions are not accessible to consciousness. For example, in an experiment Nisbett and Wilson ran, they presented two pairs of identical stockings, one on the left side and the other on the right side. The participants were asked to choose which pair of stockings they wanted. The participants chose the right most pair of stockings in a ratio of four to one. Moreover, when asked for the reasons why they chose the right most stocking, none of them said anything about the position of the stockings. When they were specifically asked if they chose the right most stocking because of its position, “virtually all subjects denied it, usually with a worried glance at the interviewer” (Nisbett & Wilson, 1977, p. 244). The minimal difference of location of otherwise identical objects should have no bearing on the choices the people make, but it did. Moreover, the factors that caused people to prefer the stocking on the right appeared not to be consciously accessible. One could draw an analogy with intuition. If factors that

influence intuitions are not accessible to consciousness, then it is hard to see how one can identify the range of reliable intuitions. After all, that would require that one can put restrictions on things of which one is not aware—something that seems odd to require of people. Of course, we might be able to restrict the range of intuitions by some conscious features. For example, we could restrict the range by a particular subject matter. But we would need *some* evidence that even in the restricted domains one is not influenced by extraneous factors that are not consciously available. That is a question that needs to be answered empirically and cannot be answered from the armchair.

4. Moderate Pessimism about Intuitions

Many philosophers have defended the use of intuitions. Two favorite apologetic moves are (a) to claim that denying intuitions have evidential weight is self-defeating, or (b) to claim that those who deny intuitions have evidential weight are offering arguments similar to those offered by a radical skeptic. These apologists paint the following picture of the critic. The critic starts with the observation that in some situations, intuitions can be led systematically astray. Because intuitions sometimes get things systematically wrong, the critic makes the inference that *no intuitions* have evidential weight. That is, because we know that intuitions sometimes get things wrong, we should not have confidence that intuitions ever get things right. Given this portrayal of the critic, the apologist normally responds in one of two ways. First, the critic's arguments are self-defeating because at one point an intuition must be invoked. Because the critic cannot make use intuitions for evidence, she cannot justify her own claim that intuitions lack evidential weight; thus, her denial that intuitions have evidential weight is self-defeating. Alternatively, the critic could embrace the self-defeating nature of her arguments. If the critic takes this alternative, then she is no different in principle than a radical skeptic. Because most do not take the radical skeptic all that seriously, we have no reason to take the critic about intuitions seriously.

George Bealer (1992, 1996, 1998) contends that the critic cannot deny intuitions have evidential weight because they are basic sources of evidence. Intuition is a basic source of evidence in the same way that our sensory faculties provide us with basic sources of knowledge. What underwrites these sources being basic is that, intuitively, there is nothing *further* that underwrites why they are sources of evidence. Because intuitions are on a par with other sources of evidence, Bealer thinks what is required of the critic of intuition is that one generate a theory that does not include intuition as a basic source of evidence. If that theory deems intuitions to be

unreliable, then we can exclude intuitions as basic sources of evidence. But, Bealer contends, that is impossible because the only way to show that intuitions are unreliable is to use intuition (1996, p. 126). Hence, to deny intuitions have evidential weight is to say that intuitions have evidential weight—an obviously self-defeating position.

Likewise, Ernest Sosa rightly points out that “it is hard to avoid appeal to direct intuition sooner or later” (Sosa, in press). When thinking about “those who reject philosophical intuitions as useless” (2007, p. 59), Sosa thinks that we merely need to reflect on how widespread and accurate appeal to intuition actually is. He thinks, “we surely do and must allow a role for intuition in simple arithmetic and geometry, but not only there. Indeed, I ask you to consider how extensively we rely on intuition. I myself believe that intuition is ubiquitous across the vast body of anyone's knowledge” (2007, p. 60). Because appealing to intuition is so widespread across our body of knowledge, “there just seems no sufficient reason for denying ourselves similar intuitive access to the simple facts involved in our hypothetical philosophical examples” (2007, p. 60). We rely on intuition in a wide variety of areas; “by parity of reasoning, therefore, it would be an overreaction to dismiss intuition just because it misleads us systematically in certain known circumstances” (1998, p. 265). If we deny that intuitions are reliable, then we “face the problem of noncircular calibration, and also a problem of a vicious regress” (2007, p. 61).¹⁵

Laurence Bonjour echoes a similar sentiment when he claims “that the practice of even those who most explicitly reject the idea of substantive a priori justification inevitably involves tacit appeal to insights and modes of reasoning that can only be understood as a priori in character, if they are justified at all” (1998, xi). When the critic of intuitions claims that intuitions are not justified, at some point the critic must appeal to intuitions—a move that seems impermissible given their criticism. That is, “no account of justification of the main radical empiricist thesis that is not in direct conflict with its truth seems to be possible” (Bonjour 1998, p. 63). Bonjour thinks “the repudiation of all a priori justification is apparently tantamount to the repudiation of argument or reasoning generally, thus amounting in effect to intellectual suicide”

15 The problem of noncircular calibration and vicious regress is, as Sosa explains, that “we must tell with priority that the source of belief B is source S1, and we must tell with priority that the source of *this* meta-belief is source S2, and with priority that the source of this meta-meta-belief is source S3, etc. So it seems out of the question to make any such priority requirement in general for our basic sources and their operation in particular instances” (200, p. 61).

(Bonjour, 1998, 5). Harvey Siegel agrees that it, “is worth re-emphasizing that in one respect the naturalized epistemologist's position is self-defeating. For it seeks to justify naturalized epistemology in precisely the way in which, according to it, justification cannot be had” (1984, p. 675). And Mark Kaplan thinks that “the naturalist's attempt to show the errors of aprioristic methodology depends for its success on consulting, and finding naturalist arguments in accord with, the very sorts of armchair intuitions whose advice the naturalists would have us ignore” (1994, p. 360).

Some philosophers think the psychological approach that some critics take ultimately puts those critics in a self-defeating position. For example, Michael DePaul writes:

[C]ritics might cite studies by cognitive psychologists showing that our intuitive judgments about the area in question are unreliable, and press the inquirer not to allow these intuitive judgments any weight at all in determining the theory she ends up accepting... But suppose the inquirer is not ignorant of the psychological studies and that she has already incorporated her belief in the results of these studies into her system of her beliefs in a way that does not require her to give her intuitive judgments no weight in her deliberation. (1998, p. 304)

But, if the critic of intuitions is right and we should not have any faith in our reflective, intuitive judgments, then:

She would have to abandon the results of her own reflection, give up, at least in part, on thinking for herself, and simply knuckle even though she firmly believes, after careful reflection on all the relevant considerations, that doing so will lead her away from truth and into error. (DePaul, 1998, p. 304)

DePaul thinks if we take the critic seriously we would have to give up what we reflectively think is correct.

Timothy Williamson argues that the critic about intuitions is self-defeating in a slightly different way. Many critics think intuitions are *merely* psychological items and nothing more, and some philosophers take the *only* source of evidence for philosophical claims to be intuitions (Williamson, in press). Williamson claims that these critics think “not that our evidence consists of the many non-psychological putative facts which are the contents of those intuitions, but that it consists of the psychological facts to the effect that we have intuitions with those contents, true or false. On such a view, evidence in philosophy amounts only to psychological facts about ourselves” (in press). Given that intuitions are psychological entities which many of us share, a

theory that respects those intuitions is better than a theory that does not respect those intuitions. But if the critic takes the psychological turn toward intuitions, then she again finds herself in a self-defeating position.

That our evidence in philosophy consists of facts about intuitions and that explanations of those facts on which intuitions come out true are better (*ceteris paribus*) than explanations on which they do not are themselves epistemological rather than psychological claims. Taken far enough, the psychologization of philosophical method becomes self-defeating. Psychologism is no more a psychological theory than the Pythagorean doctrine that everything consists of numbers is a mathematical theory. (Williamson, *in press*)

Hence, if the critic takes intuitions to be nothing more than psychological items, then the critic is in a self-defeating position.

However, skepticism about intuitions need not lead to self-defeat. As Williamson notes: We may suspect judgment scepticism is a bomb which, if it detonates properly, will blow up the bombers and those whom they hope to promote together with everyone else. But it does not follow that we can dismiss judgment scepticism as self-defeating. That the revolutionary movement would be incapable of establishing a stable new government of its own does not show that it is incapable of bringing the old government down. (*in press*)

However, Williamson continues, the argument that the critic offers cannot be legitimately confined to just intuitions. According to Williamson, the critic calls into question the following principle: “one should be confident that P (on the basis of common sense) only if its appearing (to common sense) that P is good evidence that P” (*in press*). But if the critic succeeds in criticizing that principle, then all types of appearances are called into question, some of which the critic allows. As Williamson wonders, “how can such sceptics prevent their scepticism from spreading as far as the sciences themselves? For it infects standard perceptual judgments, on which the natural sciences systematically depend” (*in press*). Hence, because the naturalist cannot restrict their criticisms to just intuitions, their position leads to radical skepticism.

Michael Bishop and I (Feltz & Bishop, *in press*) have argued that all these worries about an empirical investigation into the reliability of intuitions are worries about straw men. As we see it, one can be pessimistic about the reliability of intuitions without at the same time being a radical skeptic who calls into doubt all intuitions. That would saddle the critic with a position that is obviously false. Most of us use intuitions a lot, and those intuitions serve a vital and

productive role in our lives. After all, most of the intuitions that we use on a daily basis do a fairly good job of getting us around in the world. So, it is not the case that all of our intuitions are drastically wrong most of the time. As we write:

The moderate pessimist is not arguing that we should abandon all of our [philosophical] intuitions about everything (which would lead to [philosophical] suicide). Nor is she arguing that we can construct [philosophical] theories without ever relying on any [philosophical] intuitions. Instead, the moderate pessimist merely holds that certain sorts of [philosophical] theories....cannot be supported entirely (or perhaps even primarily) by [philosophical] intuitions. (Feltz & Bishop, in press)

Hence, the moderate pessimist is not criticizing intuitions altogether or at once. Rather, the moderate pessimist only questions some intuitions in the service of some projects. For example, the moderate pessimist may question the use of intuitions in the service of philosophical conceptual analysis but not in the service of forming immediate judgments about whether a lover is lying to them. Therefore, the moderate pessimist about intuitions can reject and question the philosophical use of intuitions while at the same time accepting the use of intuitions for other uses.¹⁶

What has been offered in this chapter is in line with the moderate pessimism expressed by Feltz and Bishop (in press). I do not claim that all intuitions in all circumstances are unreliable. Rather, I argue that there is *some* evidence to think that all intuitions in some domains are unreliable. In particular, we have reason to worry that all intuitions in some philosophically important fields are unreliable because we have not had the opportunity to calibrate those intuition by independent sources of evidence (like we have with intuitions in other, less contentious domains). Because we have some reason to think that all intuitions in some philosophically important domains are unreliable, we should make an attempt to understand, in an empirically informed way, if intuitions are reliable in those domains. Prior to investigation, I remain neutral about the actual reliability of intuitions in those domains. Given my neutrality and because the reliability of intuitions in one domain says nothing about the reliability of intuitions in another domain, I avoid radical skepticism about intuitions.

5. Conclusion

¹⁶ The moderate pessimist about intuitions can be pessimistic about intuitions in some philosophically relevant domains and not in others. For example, the moderate pessimist may reject the use of intuitions in epistemology and accept their use in logic. The main point is that the moderate pessimist about intuitions may reject the use of intuitions in all philosophically relevant domains but she need not; hence, she need not be a radical skeptic about intuition.

If the arguments in this chapter are correct, then there is presently no convincing general a priori defense of the reliability of intuitions. In fact, there is some evidence that all intuitions in certain philosophically domains are not reliable. The fact that there are some domains where there is some evidence suggesting that all intuitions are not reliable raises the question if all intuitions in many philosophically relevant domains are reliable. This question is not something that can be answered a priori. In order to determine if intuitions are in fact reliable, empirical research needs to be done to determine where intuitions are or are not reliable. An initial attempt at discovering where intuitions are reliable is the task set for the following chapters.

CHAPTER 3: EPISTEMOLOGY AND EMPIRICAL DATA

A large part of contemporary epistemology attempts to describe under what conditions one knows or is justified in having a belief. Some think an adequacy constraint on these theories is that they accord with what we think knowledge and justification are.¹⁷ That is, the norms that govern when we know or are justified in having a belief are determined by when *we* think one knows or is justified. Hence, the norms of epistemology are generated in part by us.

One way epistemic norms and analyses are said to depend on us is that they are in part generated or constrained by our intuitions about particular cases. However, there are some who think intuitions are ill suited to play this role. There have been three main ways in which intuitions in epistemology have been called into question. First, some philosophers use what they take to be commonsense intuitions to support their views. Because many believe that epistemology should capture what we think about knowledge and justification, their theories are confirmed to the extent they are supported by ordinary intuitions. However, philosophers' judgments about what intuitions the folk have are sometimes mistaken. Second, intuitions appear to vary among different cultures and socio-economic statuses. This variation might lead to relativism unpalatable to many epistemologists. Third, epistemic intuitions are subject to some biases that call into question their reliability. If intuitions are sensitive to non-epistemic features, then it is unclear for what intuitions are evidence.

There are three goals of this chapter. First, I discuss some theories that make epistemological norms in some part dependent on “us.” According to these theories, what most people think about knowledge and justification at least in part constrains and provides evidence for an adequate epistemology. Second, I point out the relevant empirical work that makes the intuition critic pessimistic that intuitions can serve this evidential role. Finally, I argue that the constructivist has several good responses to the critic. Specifically, the constructivist concedes that if philosophers are to get any theoretical mileage out of folk intuitions, they must provide the requisite empirical evidence. However, I argue that much of the evidence of intuition instability the critic provides is the result of a complex stability of intuitions. That is, there are groups of people who have stable intuitions that, when taken together, appear to indicate that the folk are

¹⁷ See Bishop and Trout (2005, p. 10) for a fuller discussion of philosophers who make reference to what we think as constraining what an acceptable epistemology is.

subject to various biases, inconsistencies, or irrationalities. This stability is argued to be sufficient to ground attenuated epistemological theory and analyses.

1. Two Ways Epistemology Depends on Us

There are two ways epistemic norms might depend on us: (1) the norms are determined by our concept of or what we mean by knowledge; (2) the norms are pragmatically determined by our desires. (1) is a typical feature of what has come to be known as Standard Analytic Epistemology. Standard analytic epistemology can be characterized as aiming “to provide a theory that captures our considered judgments about knowledge and justification” (Bishop & Trout, 2005, p. 8). According to this characterization, “if an epistemic theory forces us to radically alter our considered epistemic judgments (e.g., our epistemic judgments in reflective equilibrium), then ipso facto that theory is unacceptable” (Bishop & Trout, 2005, p. 9). Bishop and Trout call this interplay of theory and considered epistemic judgments the “stasis requirement” (2005, p. 8). For example, standard epistemology proceeds by generating theories and then testing those theories against counterexamples. These counterexamples generate intuitions. To the extent that these intuitions are not in accord with the theory, the theory is rejected. Philosophers' deeply held commitments are the standard by which theories are tested and play a critical role for acceptable epistemic theories and analyses (Bishop & Trout, 2005, p. 9-11).

One way the stasis requirement manifests itself is in the method of reflective equilibrium. Reflective equilibrium finds its first expression in Nelson Goodman (1955). Goodman states, “A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend” (1955, p. 64). Hence, reflective equilibrium is a commitment to our deeply held judgments as final arbiter of acceptable epistemic norms. Those deeply held judgments determine the acceptability of a rule or inference.

Alvin Goldman (1986) makes this commitment to common, deeply held judgments even more explicit. According to Goldman, if one is justified in believing, one must believe in accordance with a system of justification rules. These rules state whether one has a “right” to a given belief (1986, p. 59). In addition, this system of rules must be the correct rules and not just any set of rules. After all, it would seem odd that one would be justified in having a belief just by believing in accordance with some (bad) set of rules. There must be a way to determine which system of rules is correct. According to Goldman, the right system of rules is determined by a criterion of rightness (1986, p. 63). The correct criterion of rightness is determined by what

accords with our everyday intuitions. Goldman writes:

We examine what rule systems would likely be generated by each candidate criterion. We reflect on implications of these rule systems for particular judgments of justifiedness and unjustifiedness. We then see whether these judgments accord with our pretheoretic intuitions. A criterion is supported to the extent that implied judgments accord with such intuitions, and weakened to the extent that they are not. (Goldman, 1986, p. 66)

Hence, the right set of justification rules is the one that accords, more or less, with our pretheoretical intuitions about specific cases. Intuitions are guides to the correct criterion of rightness, and therefore our ordinary intuitions about knowledge in part govern what is an acceptable epistemology.

The pragmatic approach mentioned in (2) is a more recent development. It has come about largely because of dissatisfaction with Standard Analytic Epistemology. For example, Stephen Stich thinks standard analytic epistemology has failed to offer a way to decide on a criterion of rightness. According to Stich, there might be groups of people with different epistemic intuitions, concepts, or standards of knowledge. One group may intuitively accept one criterion of rightness, and another group may intuitively accept a different one. The problem is that reconciling those two criteria is impossible because there are no intuitions to determine a “meta” criterion of rightness. Because there is no way to determine the single criterion of rightness, Stich thinks that the correct standards of epistemic evaluation are ones that get us most of what we intrinsically value (which is not true belief) (Stich, 1990).

Hilary Kornblith (2002) agrees that standard analytic epistemology has failed, but he thinks that there is one thing all reasoners care about—true beliefs.

Precisely because our cognitive systems are required to perform evaluations relative to our many concerns, and to perform these evaluations accurately, the standards by which we evaluate these cognitive systems themselves must remain insulated from most of what we intrinsically value, whatever we may value. This provides a reason to care about the truth whatever we may otherwise care about. (Kornblith, 2002, p. 158)

Hence, while these philosophers do not think that epistemic evaluation depends on what we mean by knowledge or our concept of it, epistemic evaluation does rely on us in the sense that at least some of our desires figure critically into correct epistemological advice.

There are other ways in which philosophers have argued that folk intuitions are relevant to epistemological debates. Jason Stanley (2005) argues that, in addition to “traditional factors,”

our ordinary practice of knowledge ascription is sensitive to the *practical facts* of a subject's situation. The practical facts of a situation are “facts about the costs of being right or wrong about one's beliefs” (Stanley 2005, p. 6). Stanley appeals to our ordinary practices of knowledge ascription to reveal this connection. For example, when the costs of S having a false belief that p are low, Stanley contends that people are likely to agree that S knows p. However, on the basis of the same evidence, when the costs of S falsely believing p are high, he claims that people are likely to think that S does not know p. “Ordinary assertions of knowledge are made on such a basis that we can envisage someone [who possesses knowledge] in a higher-stakes situation (often a much higher-stakes situation), whom we would not think of as possessing that knowledge, given similar evidence” (2005, p. 8). An essential part of his case is that our ordinary practices of knowledge ascription reveal a connection between practical facts and knowledge. Hence, epistemic intuitions reveal to us something important about epistemology for which the best theories should be able to account.

As we can see, there is a diverse set of philosophers who make reference to our intuitions or desires to determine what epistemic norms are acceptable. What our intuitions are is an empirical matter, so what does the empirical evidence tell us about our epistemic intuitions? Exploring that question is the task of the next three sections.

2. Standard Analytic Epistemology and Empirical Data

Critics of epistemic intuitions interpret the data as indicating that intuitions are ill-suited for the role Standard Analytic Epistemology has given them. For example, it has been argued that the stasis requirement leads to an unacceptable relativism (Bishop & Trout, 2005; Weinberg, Nichols, & Stich, 2001). As we saw in Chapter 1, WNS suggest that the epistemological intuitions of philosophers could be the intuitions of a small, idiosyncratic sub-set of the entire population. Their studies directly target using intuitions about key cases as reasons for thinking that there is one true epistemological theory.

The results of WNS are especially troubling if epistemic norms are in part dependent on us. First, if epistemic norms are determined by what we mean by knowledge or by our concept of knowledge, then it looks like a plurality of epistemic norms will be legitimate. If the epistemic norms depend on our intuitions and those intuitions systematically vary between groups, then there could be more than one set of epistemic norms accepted as legitimate. Because those norms fit well with different sets of intuitions, there will be no way to decide between those epistemological theories. That leads to relativism about epistemic norms with no single, true

epistemological theory.

Not only is there systematic diversity of epistemic intuitions, some epistemic intuitions also appear to be influenced by extraneous factors. Stacey Swain, Joshua Alexander, and Jonathan Weinberg (SAW) (2008) have found that epistemic intuitions are influenced by what has been recently considered. In their study, they gave the same participant four scenarios describing a person in four different epistemic situations. (1) A clear case of knowledge where Karen, a chemist, reads that mixing two chemicals will create a toxic gas, and she forms a belief that mixing two chemicals together will create a noxious gas. (2) A clear case of non-knowledge describing Dave who gets a special feeling about what side of a fair coin will come up and forms a belief to that effect. The two other cases are contentious and are drawn from the epistemology literature. (3) The case in which SAW are particularly interested is normally taken to be a counterexample to reliabilism. This case describes Charles who gets hit in the head and subsequently almost always forms a true belief about the temperature. (4) A case describing Suzy who is traveling through a countryside with barn facades and happens to look upon the only real barn. She forms the belief that she is looking at a real barn.

SAW gave these cases to participants in various orders. They found that the order of presentation influenced the intuitions participants had about Charles. When a clear case of knowledge was presented before Charles's case, participants did not think that Charles knows what the temperature is. However, when a clear case of non-knowledge is presented before Charles's case, people thought that Charles knows. Hence, the order of presentation reversed people's intuitions about Charles.

According to SAW, this should be troubling for epistemologists who use epistemic intuitions as evidence. In the cases SAW use, the intuitions are influenced by a non-evidential factor—the order of presentation. If intuitions are influenced by non-evidential factors, then intuitions are sensitive to features that are not epistemically relevant. Because intuitions can be influenced by these non-epistemic features, we should not be confident that our intuitions accurately track what knowledge is. Hence, without evidence that our epistemic intuitions track only epistemic features, we should be especially cautious, and perhaps dispense with, using intuitions in epistemology all together.¹⁸

¹⁸ One might think that contextualism can handle these order effects. However, according to SAW, it is difficult to determine what the relevant context is. Of course, it is possible that contextualism *can* handle these order effects, one would like to have some empirical evidence that contextualism *does* handle these order effects.

3. Describing Folk Intuitions¹⁹

Whether the normal practice of knowledge ascription is sensitive to practical facts as Stanley suggests is an empirical question. Stanley presents a series of cases where the practical facts are supposed to play a role. These cases describe Sarah and Hannah driving past the bank on a Friday afternoon. They wonder if the bank will be open on Saturday so they can deposit their paychecks. In *Low Stakes*, they have no impending bills, so there is no pressure for them to deposit their checks by Saturday. Hannah says “I know the bank will be open.” In *High Stakes*, Sarah and Hannah have a bill coming due and little money in their account, so it is important that they deposit their paychecks by Saturday. Hannah states, “I don't know the bank will be open.” In both of these cases, the bank is open on Saturday. Stanley thinks, “In *Low Stakes*, our reaction is that Hannah is right; her utterance of 'I know the bank will be open' is true. In *High Stakes*, our reaction is that Hannah is also right. Her utterance of 'I don't know that the bank will be open' is true” (2005, p. 5).

In both *High Stakes* and *Low Stakes*, Sarah and Hannah are aware of the practical facts of the situation. However, Stanley thinks this is not necessary for practical facts to have an effect on our ordinary practice of knowledge ascription. *Ignorant High Stakes* is like *High Stakes*, but Sarah and Hannah aren't aware of their impending bill. Again, because the stakes are high, Stanley predicts that people are likely to say that Hannah does not know. According to Stanley, the practical facts of a situation also play a role when we assess whether a person has correctly attributed knowledge to another person. In *Low Attributer-High Subject Stakes*, Jill says that Hannah knows that the bank will be open on Saturday. Nothing is at stake for Jill, but she is unaware that Sarah and Hannah have a bill coming due. Stanley predicts that people are likely to say that Jill is wrong that Hannah knows that the bank will be open (2005, p. 5-6) (See Appendix B).²⁰

Stanley's predictions are summarized in Table 1:

19 Parts of this section appeared in Feltz and Zarpentine (manuscript).

20 Stanley describes a fifth case. In *High Attributer-Low Subject Stakes*, Hannah and Sarah have a bill coming due, and Hannah asserts that her friend Bill does not know that the bank will be open on Saturday. In this case, Stanley predicts that people are likely to agree that Hannah correctly thinks that Bill does not know. We did not give participants High Attributer-Low Subjects stakes because Stanley thinks the folk judgments about this case are mistaken. The justification for not including this scenario is that it would not be directly relevant to Stanley's thesis. That is, if it turns out that the folk do not make these judgments, then he would not have to explain away their judgments. However, if he is right, then he has an explanation for why they made those judgments. Hence, discovering what the folk think about that case would have no bearing on Stanley's central thesis that practical interests play a role in knowledge attributions.

Table 1: Folk Intuitions Predicted by Stanley

| Scenario | The Predicted Response |
|------------------------------------|---|
| Low Stakes | The person described does know. |
| High Stakes | The person described does not know. |
| Ignorant High Stakes | The person described does not know. |
| Low Attributer-High Subject Stakes | The person to whom knowledge is attributed does not know. |

Stanley thinks that these “examples involve ordinary knowledge ascriptions” (2005, p. 32).

Therefore, we can test Stanley's predictions by giving these scenarios to ordinary folk. Chris Zarpentine and I did exactly that.

We gave Stanley's cases to 152 volunteers in introductory-level philosophy classes at Florida State University.²¹ Each participant received one of the following four cases: Low Stakes, High Stakes²², Ignorant High Stakes, and Low Attributer-High Subject Stakes (see Appendix C). We asked them to rate the degree to which they agree with a statement regarding a knowledge claim made by one of the people described in the scenario (1 = 'Strongly Agree', 4 = 'Neutral', 7 = 'Strongly Disagree').²³

Table 2 illustrates the percentages of people who agreed to some extent (answered 1-3), disagreed to some extent (answered 5-7), or were neutral (answered 4) about whether the person described in the scenario knows.

21 All experiments reported in this dissertation conformed to the Florida State University Institutional Review board standards. See Appendix A for the Florida State University Institutional Review Board approval letter and sample informed consent form.

22 We reverse scored this scenario because the prompt asked the participants to indicate if it is true that Hannah *does not* know. We did not change this case because we wanted to reproduce Stanley's scenarios exactly as he presents them.

23 In all the cases we used to test Stanley's prediction, we asked participants' level of agreement to a statement of the form “Subject S know that P.” The only statement that was in a different form was for High Stakes, which took the form “Subject S does not know that P.”

Table 2: Results of Stanley's Scenarios

| | Agree to some extent | Disagree to some extent | Neutral |
|---|----------------------|-------------------------|---------|
| Low Stakes (N=34) | 56% | 41% | 3% |
| High Stakes (N=39) | 36% | 43.5% | 20.5% |
| Low Attributer-High Subject Stakes (N=40) | 27.5% | 57.5% | 15% |
| Ignorant High Stakes (N=39) | 48% | 39% | 13% |

There is no statistically significant difference between Low Stakes and High Stakes, although there is a numerical shift in the predicted direction,²⁴ and there is no statistically significant difference between Low Stakes and Ignorant High Stakes.²⁵ However, we did find a statistically significant difference between Low Attributer-High Subject Stakes and Low Stakes,²⁶ and we found a statistically significant difference between Ignorant High Stakes and Low Attributer-High Subject Stakes.²⁷

Our results do not demonstrate the pattern of ordinary knowledge ascriptions predicted by Stanley. He predicts that most people will agree that Hannah knows in Low Stakes and that in the other three cases most people will disagree that the person in the scenario knows. But in High Stakes, 43.5% disagree that Hannah knows and in Ignorant High Stakes 39% disagree. Therefore, while most people have the predicted intuitions about Low Stakes and Low Attributer-High Subject stakes, we do not find evidence for the overall pattern of intuitions predicted by Stanley.

In light of these results, Stanley might maintain that it is not critical that most people have this pattern of intuitions. All that he requires is that practical facts have *some* effect on ordinary ascriptions of knowledge. We did find a statistically significant difference between Low Stakes

²⁴ High Stakes ($M=4.2564$) and Low Stakes ($M=3.6765$), $t(71) = 1.213$, $p > 0.05$.

²⁵ Ignorant High Stakes ($M=3.589$) $t(71) = .194$, $p > 0.05$.

²⁶ Low Attributer-High Subject Stakes ($M = 4.75$), $t(72) = 2.417$, $p < 0.05$.

²⁷ $t(77) = 2.719$, $p < 0.05$.

and Low Attributer-High Subjects Stakes. Also, there is a numerical shift in the predicted direction between Low Stakes and High Stakes. These results offer support for the thesis that practical facts *do* have *some* effect on ordinary ascriptions of knowledge.

We (Feltz & Zarpentine, manuscript) argued that an alternative pair of hypotheses explain these results much better than the practical facts. First, the difference in people's intuitions about High Stakes and Low Stakes is an artifact of how Stanley describes the cases. In particular, the statement “Sarah points out that banks do change their hours,” appears in High Stakes but not in Low Stakes. This statement makes salient a possible justification defeater which might account for why more people ascribe knowledge in Low Stakes than in High Stakes. Second, the difference in people's responses between Low Stakes and Low Attributer-High Subject Stakes is the result of an *attributer effect*: people are more reluctant to agree with third-person knowledge attributions than first-person attributions.

Our results provided three reasons which support the attributer effect. First, the only significant differences we found were between third-person and first-person cases: (a) between Low Attributer-High Subject Stakes and Low Stakes, and (b) between Low Attributer-High Subject Stakes and Ignorant High Stakes. Second, the practical facts of Low Attributer-High Subject Stakes and Ignorant High Stakes are the same. If only the practical facts and traditional epistemic features influence the results in these cases, then on Stanley's view there should be no significant difference between these two cases. But there is. A third reason which supports the attributer effect is that there is a significant difference between Low Stakes and Low Attributer-High Subject Stakes but not between Low Stakes and High Stakes.

We ran a follow-up study to test our two hypotheses. First, we developed a pair of cases without the potential confounding factors that appeared in Stanley's original cases. The only difference between our cases is the stakes. If ordinary knowledge ascriptions are sensitive to the practical facts of a situation, then we should observe a difference between these cases. Second, we tested our attributer effect hypothesis by developing a case which involves third-person knowledge attribution.

In the follow-up study, we used the following three scenarios:

Minimal Low Stakes: Bill, Jim, and Sarah are hiking and they come to a ravine. There is a bridge five feet over the ravine. Bill sees Sarah and Jim cross the bridge, and Bill says to Jim, “I know that the bridge is stable enough to hold my weight.”

Minimal High Stakes: Bill, Jim, and Sarah are hiking and they come to a ravine. There is

a bridge one hundred feet over the ravine. Bill sees Sarah and Jim cross the bridge, and Bill says to Jim, “I know that the bridge is stable enough to hold my weight.”

Attributer: Bill, Jim, and Sarah are hiking and they come to a ravine. There is a bridge five feet over the ravine. Bill sees Jim and Sarah cross the bridge, and Jim says to Sarah, “Bill knows that the bridge is stable enough to hold his weight.”

The only difference between *Minimal High Stakes* and *Minimal Low Stakes* is the height of the bridge. If the results of our test of Stanley's original cases were the result of confounding factors, then we should find no difference in people's attributions in the minimal cases. Likewise, if the statistically significant results we found in our test of Stanley's original cases were due to the attributer effect, then we should find a significant difference between cases describing first-person and third-person knowledge attributions.

In our follow-up study, 119 students in introductory-level philosophy courses at Florida State University volunteered to participate in the experiment. Each participant was given one of either Minimal High Stakes, Minimal Low Stakes, or Attributer (see Appendix D). They were asked to rate the degree to which they agree with a statement regarding a knowledge claim made by one of the people described in the scenario (1 = 'Strongly agree', 4 = 'Neutral', 7 = 'Strongly disagree'). Table 3 illustrates the percentages of people who agreed to some extent (answered 1-3), disagreed to some extent (answered 5-7), or were neutral (answered 4) about whether the person described in the scenario knows.

Table 3: Results of the Minimal Cases

| | Agree to some extent | Disagree to some extent | Neutral |
|----------------------------|----------------------|-------------------------|---------|
| Minimal Low Stakes (N=41) | 54% | 29% | 17% |
| Minimal High Stakes (N=39) | 56% | 18% | 26% |
| Attributer (N=39) | 33% | 31% | 36% |

These results clarified our first experiment. In the simplified cases, the only relevant difference between Minimal High Stakes and Minimal Low Stakes is the practical facts. Stanley's view predicts that people should be less likely to attribute knowledge to Bill in Minimal High Stakes than in Minimal Low Stakes. But we do not find this. In fact, there is no statistically significant difference in how people ascribed knowledge in Minimal High Stakes and Minimal Low Stakes.²⁸ These results suggest that practical facts *do not* play a role in ordinary knowledge attribution. These results also supported our attributer effect hypothesis. Considerably fewer people agree that the person knows in the Attributer case (33%) compared to the percentage of people who think that the person knows in Minimal Low Stakes (54%) or Minimal High Stakes (56%).²⁹ Thus, we have good reason to think that what explains the results of the Low Attributer-High Subject Stakes in our first study is the *attributer effect* and *not* the practical facts.

One worry with our follow-up study might be that there is no room for the practical facts to play a role. Bridges are normally constructed with iron, steel, concrete, or heavy timber—all of which are sufficient to hold a normal person's weight. If our participants assume that the bridges are stable, then the difference in practical facts between our minimal cases would not affect ordinary knowledge ascriptions regardless of how high the bridge is. If this is the case, then Stanley would not *expect* these cases to provide evidence for his thesis.³⁰

To address this worry, we ran a third experiment. We minimized possible confounding factors in Stanley's original cases by creating the following two scenarios:

Simplified High Stakes: Hannah and her sister Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, it is very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. Hannah says to Sarah, 'I know that the bank will be open tomorrow'.

Simplified Low Stakes: Hannah and her sister Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks.

28 Minimal High Stakes ($M = 3.2308$) and Minimal Low Stakes ($M = 3.2927$), $t(78) = .165$, $p > 0.05$.

29 There is a statistically significant difference between Minimal High Stakes and Attributer, $t(76) = 2.062$, $p < 0.05$. There is a near significant difference between Minimal Low Stakes and Attributer, $t(78) = 1.739$, $p = .086$. And, when we combine the results of Minimal High Stakes and Minimal Low Stakes, we find a statistically significant result between the Attributer and the minimal non-attributer cases, $t(117) = 2.062$, $p < 0.05$.

30 We would like to thank Jason Stanley (personal communication) for bringing this worry to our attention.

Since they do not have an impending bill coming due, it is not very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. Hannah says to Sarah, 'I know that the bank will be open tomorrow'.

In *Simplified Low Stakes* and *Simplified High Stakes*, the only difference is the practical facts of the situations. In *Simplified High Stakes*, it is very important for Hannah to deposit her check because she has a bill coming due. In *Simplified Low Stakes*, with no bill coming due, it is not very important for Hannah to deposit her check. In Stanley's original cases, "our intuition in *High Stakes* [is] that Hannah does not know that the bank will be open" (Stanley 2005, p. 129) and, "In *Low Stakes*, our reaction is that Hannah is right; her utterance of 'I know the bank will be open' is true" (Stanley, 2005, p. 5). Because the simplified cases are versions of Stanley's original cases without the possible confounding factors, most people should think that Hannah knows in *Simplified Low Stakes* but does not know in *Simplified High Stakes* if Stanley is right.

Eighty-three undergraduates in introductory-level philosophy classes at Florida State University volunteered to participate in the experiment. Each participant was given one of either *Simplified High Stakes* or *Simplified Low Stakes* (see Appendix E). They were asked to rate the degree to which they agree with a statement regarding a knowledge claim made by one of the people described in the scenario (1 = 'Strongly agree', 4 = 'Neutral', 7 = 'Strongly disagree'). Table 4 illustrates the percentages of people who agreed to some extent (answered 1-3), disagreed to some extent (answered 5-7), or were neutral (answered 4) about whether the person described in the scenario knows.

Table 4: Results of the Simplified Cases

| | Agree to some extent | Disagree to some extent | Neutral |
|-------------------------------|----------------------|-------------------------|---------|
| Simplified Low Stakes (N=41) | 44% | 34% | 22% |
| Simplified High Stakes (N=42) | 43% | 38% | 19% |

If Stanley is right, we should expect that people judge that Hannah knows in *Simplified Low Stakes*, but does not know in *Simplified High Stakes*. This pattern of responses is not found.

Again, there is no statistically significant difference in responses.³¹ These results confirm the results of our second experiment suggesting that practical facts do not play a role in people's ordinary practice of knowledge ascriptions.

But perhaps our revision of Stanley's original scenarios still does not make the practical facts salient. It is easy to miss the cost of Hannah having a false belief in Simplified High Stakes. If the practical facts are not salient to the participants, then again we should not expect to find significant differences in participants' responses. To make the practical facts more salient, we ran another experiment using cases suggested by Jason Stanley (personal communication):³²

High Stakes Bridge: John is driving a truck along a dirt road in a caravan of trucks. He comes across what looks like a rickety wooden bridge over a yawning thousand foot drop. He radios ahead to find out whether other trucks have made it safely over. He is told that all 15 trucks in the caravan made it over without a problem. John reasons that if they made it over, he will make it over as well. So, he thinks to himself, 'I know that my truck will make it across the bridge.'

Low Stakes Bridge: John is driving a truck along a dirt road in a caravan of trucks. He comes across what looks like a rickety wooden bridge over a three foot ditch. He radios ahead to find out whether other trucks have made it safely over. He is told that all 15 trucks in the caravan made it over without a problem. John reasons that if they made it over, he will make it over as well. So, he thinks to himself, 'I know that my truck will make it across the bridge.'

The only difference between *High Stakes Bridge* and *Low Stakes Bridge* is that in the former the bridge spans a 'yawning thousand foot drop' while in the latter the bridge is over a 'three foot ditch.' Obviously, John falsely believing his truck will make it across the bridge in High Stakes Bridge is very costly—certain death—whereas the costs in Low Stakes Bridge are likely some minor injuries and embarrassment. If practical facts are to play a role in ordinary knowledge ascription, we should expect them to play a role here where the costs of having a false belief are so severe.

One hundred and forty students in lower level geography, criminal justice, and political

31 Simplified High Stakes ($M=3.8333$) and Simplified Low Stakes ($M=3.8049$), $t(78)=-.071$, $p>0.05$. It should also be noted that there is neither a significant difference between Simplified High Stakes and Minimal High Stakes, $t(79)=1.533$, $p>0.05$, nor between Simplified Low Stakes and Minimal Low Stakes $t(80)=1.325$, $p>0.05$.

32 Thanks again to Jason Stanley for pressing this worry about our third study, and for suggesting the High Stakes Bridge and Low Stakes Bridge cases.

science classes at Florida State University volunteered to participate in the experiment. Each participant was given one of either High Stakes Bridge or Low Stakes Bridge (see Appendix F). They were asked to rate the degree to which they agree with a statement regarding a knowledge claim made by one of the people described in the scenario (1 = 'Strongly agree', 4 = 'Neutral', 7 = 'Strongly disagree'). Table 5 illustrates the percentages of people who agreed to some extent (answered 1-3), disagreed to some extent (answered 5-7), or were neutral (answered 4) about whether the person described in the scenario knows.

Table 5: Results of the Bridge Scenario

| | Agree to some extent | Disagree to some extent | Neutral |
|---------------------------|----------------------|-------------------------|---------|
| Low Stakes Bridge (N=70) | 63% | 27% | 10% |
| High Stakes Bridge (N=70) | 50% | 36% | 14% |

We do not find the pattern of responses predicted by Stanley. Given the severity of John having a false belief in High Stakes Bridge, Stanley's view predicts that *significantly* more people will say that John does *not* know in High Stakes Bridge than in Low Stakes Bridge where the costs of being wrong are only minor injuries and embarrassment. In Low Stakes Bridge, 27% of participants say that John does not know. However, in High Stakes Bridge, where the cost of being wrong is a certain and horrifying death, 36% of participants disagree—a statistically insignificant difference of 9%.³³ Thus, we found no evidence that the practical facts in High Stakes Bridge have an effect on ordinary ascriptions of knowledge. This result, with the results of our previous studies, reinforces the conclusion that Stanley does not accurately describe our ordinary epistemic practices.

4. Order Effects and Individual Differences

So far, we have seen that (1) epistemic intuitions appear to vary among different cultures and socio-economic statuses; (2) people's epistemic intuitions are influenced by extraneous

³³ High Stakes Bridge ($M= 3.8286$) and Low Stakes Bridge ($M= 3.4$), $t(138) = -1.370, p > 0.05$. Of note, High Stakes Bridge does not significantly differ from either Minimal High Stakes $t(107) = 1.634, p > 0.05$ or Simplified High Stakes $t(110) = 0.013, p > 0.05$. Likewise, Low Stakes Bridge does not significantly differ from Minimal Low Stakes $t(109) = 0.312, p > 0.05$ or Simplified Low Stakes $t(109) = 1.185, p > 0.05$.

factors; and (3) philosophers are sometimes mistaken about what the folk epistemic intuitions are. One further interesting question is if the order effect reported by SAW is explainable by individual differences. Edward Cokely and I tested this hypothesis. We thought that individual differences could define groups who are at least in part responsible for some of these results. There is good reason to think this because WNS's data suggest *groups* of people have systematically different epistemic intuitions (e.g., Indians, Westerners, High v. Low SES's). It is probable that groups of people share systematically different intuitions about the cases SAW present.

We ran SAW's experiment along with collecting data about individual differences (e.g., gender, personality traits, and cognitive impulsivity). Because the only significant results SAW found were between two survey versions, we decided to give half of the participants the Karen (K), Charles (C), Dave (D), Suzy (S) order and half the SDCK order (see Appendix G). Participants were 110 volunteers in low level philosophy classes at Florida State University. Participants were asked to rate to what extent they agree that a person in the scenario knows (1 = Strongly Agree, 3 = Neutral, 5 = Strongly Disagree).

Table 6: Overall Means to SAW's Scenarios

| | KCDS | SDCK |
|---------|--------|--------|
| Charles | 3.3091 | 2.7636 |
| Karen | 2.1091 | 1.6852 |

As with the studies of SAW, we found an order effect with Charles.³⁴ However, we also found an order effect with Karen.³⁵

We thought that we would be able to identify groups with stable individual differences. That is what we found. Specifically, we found that men and women are affected differently by the order of presentation. Table 7 represents men's means, and Table 8 represents women's

³⁴ $t(108) = 2.296, p < 0.05$.

³⁵ $t(107) = 2.000, p < 0.05$. There were 109 participants in this condition because one participant left an answer blank.

means.

Table 7: Men's Mean Responses to SAW's Scenarios

| | KCDS | SDCK |
|---------|--------|--------|
| Charles | 3.4074 | 2.6944 |
| Karen | 1.7407 | 1.6571 |

Table 8: Women's Mean Responses to SAW' Scenarios

| | KCDS | SDCK |
|---------|--------|--------|
| Charles | 3.2143 | 2.8947 |
| Karen | 2.4643 | 1.7368 |

An independent samples t-test shows that the different orders influence men's judgments for Charles,³⁶ but not for Karen.³⁷ Likewise, an independent samples t-test showed that the different orders influenced women's judgments for Karen,³⁸ but not for Charles.³⁹ Importantly, there is a qualitative shift in men's intuitions about Charles such that when Charles is presented after Dave, men think that Charles knows whereas when Charles is presented after Karen, men judge that Charles does not know.

Hence, in this sample, the order effect SAW report is only significant for men and not for women. However, in addition to the order effect discovered by SAW, we found that there was an order effect on the clear case of knowledge—Karen. And, oddly, this order effect was significant for women and not for men. While there is not a qualitative shift, there was a trend toward stronger agreement when Karen is presented last in the order as opposed to first.

³⁶ $t(61) = 2.123, p < 0.05$.

³⁷ $t(60) = .743, p > 0.05$.

³⁸ $t(45) = 2.042, p < 0.05$.

³⁹ $t(45) = .927, p > 0.05$.

5. The Relevance of the Empirical Data to Epistemological Debates

5.1 Critical Implications

Three main critical implications can be drawn from the empirical data presented in this chapter. First, the evidence suggests that philosophers often misdescribe folk intuitions. For example, Stanley draws upon folk intuitions to reveal the common practice of knowledge ascription to motivate his theory. However, the intuitions which are supposed to motivate his theory simply are not the intuitions that most people have. Here the critic points out that philosophers simply cannot rely, a priori, on their own intuitions to be typical nor can they rely on their a priori determination of what folk intuitions are. One possible explanation for why philosophers get things wrong is that their theories may get in the way of their thinking about folk intuitions or they may fall prey to the false consensus effect.⁴⁰ Therefore, in order for philosophers to draw support from folk intuitions, they must provide empirical evidence that the folk have the intuitions that philosophers say they do.

The second implication is that often epistemic intuitions are biased in ways that are not obvious to the person who has them. As SAW's results suggest, people's epistemic intuitions can be influenced by what has been recently considered. If intuitions are as unstable as SAW think, then neither are they clear evidence for epistemic claims nor can they be used to base analyses of epistemic concepts. They are not clear sources of evidence because they are sensitive to more than just the epistemic features of scenarios. This causes them to fluctuate in unexpected ways. If they fluctuate, then it is unclear how we are to use those intuitions in conceptual analyses or epistemological theory—it just isn't clear *which* intuitions we should use. At a minimum, those who think intuitions are reliable need to provide some empirical evidence that they are.

Third, following WNS, we should be hesitant to think that philosophers' intuitions have special access to the truth that non-philosophers' intuitions lack. That is, philosophers' intuitions

⁴⁰ The false consensus effect is the tendency to think that the way one thinks is more prevalent in the general population than it really is.

are just as likely to be culturally local and not indicate, in any significant sense, one universal truth about knowledge and justification. As we have seen, people's intuitions can be influenced in ways of which they are not aware. There is no evidence suggesting that philosophers are any different. So, philosophers should be cautious about using their intuitions as evidence.

We see that some epistemic intuitions vary systematically from culture to culture and from social economic status to socio economic status. We also find that some epistemic folk intuitions are sensitive to non-epistemic factors. In addition, the evidence suggests that some philosophers are bad judges about what epistemic intuitions the folk have. In sum, this evidence might be sufficient for the critic to claim that epistemic intuitions should not play a role in epistemology altogether. But is this additional, stronger claim warranted?

5.2 Constructivist Implications

The strong claim that intuitions should play no role in epistemological theory is too strong. There is evidence that intuitions are stable enough to be used in epistemology, but that evidence might be surprising (and unwelcome) to some. That is, it looks like there are *groups* of people who express stable intuitions about cases. The troubling sensitivity to non-epistemic features of cases appears to be a surface phenomenon masking a deeper complexity. If there are groups of people who have stable intuitions, then their intuitions might be used to test conceptual analyses or theories in epistemology.

First, the studies of WNS are only threatening to a particular view of epistemology that assumes that a single analysis of knowledge and justification can be given that more or less matches pre-theoretical intuitions. Looked at differently, the studies of WNS suggest that there *is* stability among groups. Most Westerners think that one does not know in a Gettier case and most East Asians think that one does know in a Gettier case. Moreover, those in lower socio-economic statuses have different sets of intuitions than those of a high socio-economic status about some epistemic scenarios. These results suggest that there is some stability in those groups of people. Hence, we have some reason to think that these different groups simply have different, stable intuitions about knowledge and justification. While this may be upsetting to some, it certainly gives one hope that we can give analyses and theories employing those different intuitions. While constructivists (and philosophers) may have to give up the notion that there is only one analysis of knowledge or justification that comports with all pre-theoretical intuitions, there is no reason to think that we cannot offer analyses and theories of justification and knowledge

altogether.⁴¹

There is more evidence that there are stable groups of people with different intuitions. As will reemerge in later chapters, we find that the seeming irrational responses to some epistemic scenarios are surface phenomena. The seemingly irrational answers are the result of underlying stability in intuitions of groups of people. In the present context, the order effect on Charles that SAW report is only significant for men and not for women. Women's intuitions remained stable concerning Charles regardless of the order of presentation. Moreover, the order effect found on Karen is only significant for women and not for men. That suggests that for clear cases of knowledge, men are not affected by non-epistemic factors whereas women are. This may indicate that men and women have different concepts of knowledge. Most women may have a reliabilist conception of knowledge, whereas most men may have a justified true belief conception of knowledge. If we can identify the groups of people who are responsible for the seeming instabilities, and we find that these sub-groups' intuitions are not unstable, then we have hope that we can offer analyses of those different concepts or theories that take into account those intuitions. Those intuitions offer evidence for *their* concepts which we should be able to analyze. Likewise, we could systematize their intuitions into theories.

Of course, some might worry that this type of stability is too fragmented to be of much use for conceptual analysis or theory formation. There simply would be no single correct conceptual analyses or theory. There would be no way to determine which analysis or theory of knowledge is true. On one hand, this objection is right. The approach offered here does entail that there might be many concepts and theories. This might be objectionable to those who attempt to offer an analysis of one single concept of knowledge or one true theory. However, if we want our analyses to be constrained by commonsense intuitions, and what I have argued is right, then there simply never was one folk concept there to analyze in the first place. On the other hand, this approach does not address those who do not want their theories or analyses constrained, at least in part, by what “we” think. They can continue to do philosophy as it has traditionally been done.⁴² So, either we are open to the possibility of a plurality of concepts and theories, or we admit that theories and concepts might not be in contact with the way actual

41 This does not rule out the possibility of *one, true* analysis or that trying to construct such an analysis is not a legitimate project. Rather, philosophers engaged in such a project must either (1) not draw support from the folk, (2) explain why certain groups have systematically different intuitions, or (3) provide evidence that there is no intuition diversity.

42 Of course, they run the risk of their theories being nothing more than a “philosophical fictions” (Mele 2001, p. 27).

people think about knowledge and justification.

One might also worry that epistemology is a *normative* endeavor that is supposed to tell us under what conditions we know and are justified in having a belief. If we follow the approach offered here, then one knows and is justified just as long as one's beliefs conform to the epistemic principles that one endorses. If intuitions are in part supposed to constrain and provide evidence for adequate theories and analyses, and, as I've argued, there appears to be systematic stable epistemic intuitions, then there could be a variety of acceptable epistemic norms. That might have the odd implication that one knows or is justified just in case one satisfies the conditions that are indicated by one's intuitions about cases. That entails that one's determination of knowledge and justification is self-certifying. But that is an odd result for a normative theory that, among other things, is supposed to tell us when and how we are justified in having a belief or know. This may make one think something has gone wrong.

There are three responses to this objection. First, it might turn out that the objection is right. In order to avoid this result, one may have to give a meta-criterion of epistemic rightness. We have good reason to be suspicious that a meta-criterion is forthcoming. It may turn out that philosophers who take folk intuitions seriously simply could never give a meta-criterion because there was never one to be had. Instead, these philosophers may have been providing epistemological theories that accord with their own intuitions. So, they have been simply providing self-certifying epistemological theories and analyses of epistemic concepts all along. Second, if we adopt the approach offered here, then we can *rule out* some epistemological theories as offering notions of justification and knowledge that accord with the folk. For example, we may be able to rule out the theory offered by Stanley because it might not accord with any set of folk intuitions about knowledge and justification.⁴³ Ruling out these theories certainly is progress. Finally, the evidence offered in this chapter is only offered as an argument for the possibility of conceptual analyses and theory formation as the constructivist sees them. The larger normative issues cannot be settled here.

Thus, there is hope for the constructivist. While the constructivist agrees that philosophers can no longer rest comfortably in their studies contemplating what the commonsense conception of knowledge is, philosophers do not need to give up conceptual

43 Should appeals to intuitions be democratic? Just because most people don't share Stanley's intuitions, it does not mean that there is not some group that does. Of course, this is an empirical question. There is presently no *evidence* that folk intuitions are affected by practical facts. If there is no evidence for this effect, one may think it permissible to rule out those theories until the right kind of evidence is presented. After all, until that point one would not have reason to believe that practical facts do play a role in folk epistemic intuitions.

analyses using intuitions as evidence. Rather, the philosopher must be satisfied with the type of conceptual analyses and theory that are available. Namely, ones that take into account the plurality of concepts and ones that do not necessarily attempt to give one, monolithic analysis or theory.

6. Conclusion

In this chapter, I have shown that there are serious worries about using epistemic intuitions as evidence for epistemological theories. Folk epistemic intuitions are variable, unstable, and poorly predicted by philosophers. These results provide reasons to be cautious when appealing to intuitions about epistemological claims. However, these worries can be mitigated. As it turns out, some of the worries about epistemic intuitions are generated only by surface phenomena of a more complex stability in folk intuitions. The constructivist counsels that one should provide the requisite empirical data about intuitions. Once one does, analyzing folk concepts and systematizing folk intuitions are possible even if doing so might lead to a plurality of acceptable theories and concepts.

CHAPTER 4: ETHICS, INTUITIONS, AND EMPIRICAL DATA

In the last chapter, I argued that those who wish to use intuitions in epistemology are constrained by empirical data about intuitions. As in epistemology, one might think that ethics may be impervious to empirical facts because of the is/ought gap. Simply describing the intuitions that some people have should have no purchase on whether something is wrong or right, good or bad, permissible or impermissible. Therefore, ethical theories and analyses of moral concepts are insulated from empirical evidence about intuitions.

However, philosophers have used empirical claims about folk intuitions in two ways: (1) as evidence for their theories or analyses; (2) as adequacy constraints on acceptable ethical theories or conceptual analyses. Because folk moral intuitions are used as evidence and adequacy constraints, these two uses give what we think about moral cases a critical role in ethical theorizing and conceptual analysis.

Critics of intuitions have called into question philosophers using folk moral intuitions in two ways. First, just as with epistemology, sometimes philosophers misdescribe what moral intuitions the folk have. If philosophers do this, then those misdescriptions cannot provide evidence for or constrain ethical theories and analyses of moral concepts. What the critic argues is that we need some evidence that philosophers accurately describe folk moral intuitions. That can only be done by providing the appropriate empirical evidence.

Second, there is evidence that folk intuitions may be too unreliable to be used in ways (1) and (2). Some moral intuitions are subject to biases that undermine their reliability. If some moral intuitions are unreliable, then that calls into question using moral intuitions as evidence. For example, sometimes ethical intuitions are sensitive to features of scenarios that should be irrelevant from an ethical standpoint. If some moral intuitions are so influenced, we cannot be sure which intuitions are the result of an error or bias. We cannot be sure which intuitions are the result of an error or bias because we do not already know the right answers to many important questions in ethics. Therefore, if some moral intuitions are unreliable, it is unclear that any moral intuitions should serve as evidence, as adequacy constraints, or inform our moral concepts and ethical theories.

In this chapter, I do three things. First, I discuss some philosophical uses of folk intuitions

in the two ways mentioned above. Second, I discuss some empirical evidence that suggests intuitions are ill-suited to play the pertinent roles. Finally, I argue there are good reasons to think that the constructivist can overcome some of the critical worries. In particular, I argue that the constructivist is wise to concede that the requisite empirical data must be provided to underwrite any arguments based on folk intuitions. But the constructivist should argue that the critic has overreached. As in epistemology, the instability of ethical intuitions is the result of a deeper stability—there are groups of people who have stable moral intuitions. This stability is argued to be sufficient to ground ethical theorizing and analyses of ethical concepts.

1. Empirical Claims in Ethics

1.1 Descriptions of Moral Platitudes

Several philosophers make empirical claims in ethics. Michael Smith (1994) gives empirical claims center stage in his ethical theory. Smith thinks that there are platitudes that surround moral concepts. These platitudes are the “descriptions of the inferential and judgmental dispositions of those who have mastery of the term 'rightness'” (Smith, 1994, p. 39). These platitudes in turn inform us of the correct analysis of moral concepts. Namely, an analysis of a concept is correct “just in case it gives us knowledge of all and only the platitudes which are such that, by coming to treat those platitudes as platitudinous, we come to have mastery of that concept” (Smith, 1994, p. 31). These platitudes are defined in terms of what a “normal perceiver under standard conditions” would think are platitudes (Smith, 1994, p. 29).

Smith thinks there are at least five platitudes that surround the concept of moral rightness. The first is that moral judgments are practical—when one judges that it is right to perform action *A*, everything else being equal, then one will have some motivation to *A*. The second is that moral claims are objective. For example, when one person thinks *A* is right and another person thinks *A* is not right, at least one of them must be wrong (Smith, 1994, p. 39). Third, most people treat it as platitudinous that moral facts supervene on non-moral facts. Fourth, the substance of moral claims is the betterment of the human condition. Fifth, there is a procedure by which rational people can settle moral disputes (Smith, 1994, p. 40). These five platitudes—practicality, objectivity, supervenience, substance, and procedure—are what any adequate analysis of moral concepts must take into account (Smith, 1994, p. 41). Therefore, Smith uses folk intuitions as adequacy constraints and evidence, and those two things are dependent on what “we” think are moral platitudes.

Russ Shafer-Landau also appeals to moral platitudes. For example, he asserts that “[o]nly

cognitivism straightforwardly preserves ordinary talk of moral truths” (Shafer-Landau, 2003, p. 23). Shafer-Landau agrees with Smith that objectivity permeates normal talk. When one says that *A-ing* is right yet another another says that *A-ing* is not right, at least one of them must be wrong. Additionally, Shafer-Landau contends that “we believe that moral argument can take the logical form of other kinds of argument. We think of sentential operators in moral sentences as truth-functional” (2003, p. 23). The notion that we think moral arguments are like any other arguments is closely akin to Smith's notion that there is a procedure to settle ethical disputes. When we are faced with a disagreement in ethics, just like any other field, we use arguments. Thus, there is a shared, commonsense way to resolve ethical disputes.

There is an interesting difference between Smith and Shafer-Landau. Whereas Smith thinks that the practicality requirement is a moral platitude, Shafer-Landau thinks that it is not. One way practicality has been questioned is with the possibility of the amoralist—one who understands that some things are wrong yet has no motivation not to do those things. As Shafer-Landau notes, “sketching cases of putative amoralism can test the degree to which we are committed, prior to developed metaethical theorizing, to the possibility of the amoralist. I think that our common-sense intuitions do create (revisable, rebuttable) philosophical presumptions” that the amoralist is possible (2003, pp. 146-7). Because Smith thinks practicality is a platitude and Shafer-Landau does not, they cannot both be right. Moreover, determining whether Smith or Shafer-Landau is right is *impossible* for us to do from the armchair. The only way to determine who is right is to do some empirical leg-work.

1.2 Framing Effects on Moral Intuitions

Not only are there worries that call into question philosophers' understanding of what intuitions the folk have, there are also worries about the reliability of intuitions. As Walter Sinnott-Armstrong (in press) argues, if there are framing effects on moral intuitions, then intuitions cannot be non-inferentially justified. Sinnott-Armstrong offers several types of framing effects that may influence moral intuitions: (1) *Word* framing effects occur when different, extensionally identical descriptions are used and intuitions are changed based on those different descriptions; (2) *order* framing effects occur when the order in which scenarios are presented influences intuitions; and (3) *context* framing effects occur when the context in which something is presented alters intuitions. In all these types of framing effects, the cases are supposed to be extensionally identical. So, our intuitions should be the same regardless of frame. But sometimes they are not. Sinnott-Armstrong thinks that these framing effects show that it is possible that

framing effects exist for moral intuitions in general.

Framing effects are a particular problem for those who take intuitions to be basic, non-inferential sources of evidence. According to Sinnott-Armstrong, in order for moral arguments to avoid regresses, some of the beliefs on which those arguments are grounded cannot be grounded by further beliefs—there must be beliefs that serve as regress blockers. Some think that intuitions can serve this regress blocking role. Some philosophers contend that at least some intuitions are non-inferentially justified (Audi 2005; Tolhurst 1990, 1998). That is, just by adequately holding, seeing, not having a reason to doubt, or understanding the content of those beliefs, those beliefs are justified. These beliefs, then, can serve as regress blockers to avoid skepticism about justification.⁴⁴

The possibility of framing effects poses a potential problem for the non-inferential justification of intuitions. If our moral intuitions are subject to framing effects, and framing effects are sufficient to call into question an intuition's reliability, then it is unclear that intuitions can be adequate regress blockers. At a minimum, if intuitions are to serve as regress blockers, then they must be reliable. But there are some framing effects on moral intuitions and those framing effects call into question those intuitions' reliability. Because we know that some intuitions are not reliable, one should worry that whatever the candidate regress blocking intuition is, it might be unreliable. In order to have confidence that the proposed regress blocking intuition is reliable, one would need some evidence over and above intuitions to certify that it is reliable. Obviously, one cannot use other intuitions as non-inferential justification because the same worry would apply to those intuitions. If moral intuitions are unreliable or if we do not have good reason to think that moral intuitions are reliable, then we should have little faith that they, in and of themselves, get at moral truths or concepts. What moral intuitions need is some kind of confirmation that they are reliable that goes above and beyond the intuitions themselves—and that just is a form of inferential justification.

There are a few examples of framing effects on moral intuitions. Lewis Pertriovich and Patricia O'Neil (1996) report framing effects on intuitions about the classic trolley problem cases. When the trolley problem is framed in terms of saving five people, most people agree to throwing the switch. However, when the trolley problem is framed in terms of killing one person,

⁴⁴ Sinnott-Armstrong may have an objectionably broad notion of what intuitions are. According to Sinnott-Armstrong, intuitions are beliefs that are arrived at non-inferentially. More traditionally, intuitions are considered to be immediate intellectual seemings that are not beliefs. As I have defined intuitions in the Introduction, what Sinnott-Armstrong takes to be intuitions are allowed because they are quickly formed judgments of the sort that have played important roles in philosophy of the past fifty years.

most people disagree with throwing the switch. Because these cases employ a word framing effect and are extensionally identical, it cannot be that both sets of intuitions are correct. At least one of them must be wrong. But it is unclear how we are to decide which ones are correct solely by reference to intuitions.

Kahneman and Tversky (1984) also offer another framing effect on moral intuitions. In what has been come to be called the Asian flu case, participants are told that a rare strain of the flu is expected to kill 600 people. The government is thinking about which of two programs to implement. Participants were divided into two groups. The first group was given the following two options:

If Program A is adopted, 200 people will be saved.

If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved. (Kahneman & Tversky, 1984, p. 343).

A separate group of participants is given the following two options:

If Program C is adopted, 400 people will die.

If Program D is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die. (Kahneman & Tversky, 1984, p. 232)

Kahneman and Tversky claim that Programs A and C are logically identical, and Programs B and D are logically identical so there should not be a difference in preferences for the programs. However, they find that 72% of participants prefer Program A and 78% of participants prefer Program D—a violation of the expectation that people should not switch their preferences (Kahneman & Tversky, 1984, p. 232). The most plausible explanation is that the different ways the programs are framed (either positively or negatively) influence intuitions, which is simply an instance of a word framing effect. These two examples at least give us some reason to worry that no intuitions are non-inferentially justified.

In the face of this evidence, the intuition skeptic may want evidence that some intuitions are non-inferentially justified. Obviously, one cannot give a satisfactory response to the critic by making appeal to other moral intuitions. And, of course, this point generalizes to all moral intuitions. All moral intuitions need some independent, inferential confirmation in order for their reliability to be established. Absent that, the framing effects give us some reason to think that moral intuitions are not reliable indicators of our moral concepts.

2. Ethics and Empirical Data

2.1 Empirical Data and Objectivity

Shaun Nichols (2002) tested whether objectivity and practicality are moral platitudes. If most people are not objectivists, then both Smith and Shafer-Landau are wrong about what intuitions the folk have. In addition, the empirical data can help clarify the disagreement between Shafer-Landau and Smith about the practicality requirement.

To test whether objectivity is a moral platitude, Nichols gave the following scenario to college undergraduates:

John and Fred are members of different cultures, and they are in an argument. John says, "It's okay to hit people just because you feel like it," and Fred says, "No, it is not okay to hit people just because you feel like it." John then says, "Look you are wrong. Everyone I know agrees that it's okay to do that." Fred responds, "Oh no, you are the one who is mistaken. Everyone I know agrees that it's not okay to do that." (Nichols, 2004b, p. 9)

Participants were asked to indicate who was right. They could respond that John is right and Fred is wrong because it is okay to hit people when you feel like it, or that Fred is right and John is wrong because it is not okay to hit somebody when you feel like it, or that nobody is right because there is no fact of the matter. The last option is taken to be the non-objectivist response whereas the first two are considered to be objectivist responses. In these series of experiments, there was a control question to make sure that the responses to the moral question are not driven by a more general denial of objectivity. So, the participants were also given a world version of the Fred and Sam scenario. In the world scenario, the disagreement is about whether the earth is flat. If the participants responded that there is no fact of the matter about whether earth is flat, then their answers were not used in analyses because they expressed a more general kind of non-objectivism and not specifically moral non-objectivism.

The responses to Nichols's experiments suggest that not all people are objectivists. In the restricted set of responses that excluded the more robust kind of non-objectivism, 55 out of 148 people responded as non-objectivists (about 37%). These numbers suggest that a sizable minority of people are non-objectivists. If we speculate that the people who are non-objectivists about the earth also responded non-objectivist to the moral question ($N=37$), the percentage jumps to 50%. If these figures are right, then those who think the folk are objectivists about ethics are either wrong or must at least explain why 37% (and possibly 50%) of people who responded as non-objectivists have mistaken intuitions. In any event, these data suggest that not all people think

that objectivity is a moral platitude.

2.2 Empirical Data and Practicality

Another way empirical data are relevant to theoretical claims in ethics is that they can help clarify disputes. As mentioned, Smith and Shafer-Landau disagree about whether the folk endorse the practicality of moral judgments. In an attempt to discover whether the folk endorse the practicality requirement, Nichols used the following probes:

Psychopath: John is a psychopathic criminal. He is an adult of normal intelligence, but he has no emotional reaction to hurting other people. John has hurt, and indeed killed, other people when he has wanted to steal their money. He says that he knows that hurting others is wrong, but that he just doesn't care if he does things that are wrong. Does John really understand that hurting others is morally wrong?

Mathematician: Bill is a mathematician. He is an adult of normal intelligence, but he has no emotional reaction to hurting other people. Nonetheless, Bill never hurts other people simply because he thinks that it is irrational to hurt others. He thinks that any rational person would be like him and not hurt other people. Does Bill really understand that hurting others is morally wrong? (Nichols, 2002, p. 289)

Surprisingly, 86% of the people surveyed thought that the psychopath truly understands that hurting other people is wrong even though he has no motivation not to do so. Another surprise is that “a majority of subjects denied that the mathematician really understood hurting others is morally wrong” (Nichols, 2002, p. 289). This is contrary to what we should expect if the practicality requirement is a moral platitude. These results suggest that it is not a platitude that people think those who judge it right to *A* also have motivation to *A*.⁴⁵

2.3 Worrisome Framing Effects

Empirical data are also relevant to ethical debates because they can call into question the reliability of moral intuitions. Thomas Nadelhoffer and I (Nadelhoffer and Feltz, manuscript) ran a set of experiments inspired by Sinnott-Armstrong's arguments. We thought there would be a difference in intuitions about the permissibility of throwing the switch in the classic trolley problem depending on how the scenario is framed. So, we ran a study to test that hypothesis. In light of the research on attributional biases, we wanted to see whether people's moral intuitions and judgments concerning the trolley problem differed depending on whether participants were

⁴⁵ It is consistent that Bill does have some motivation not to hurt other people. That is, there could be some not emotional motivation. But, the Mathematician case is interesting because it does seem to bar an important and common source of motivation for the proponent of the practicality requirement.

actors or observers.⁴⁶ Participants were 85 undergraduate volunteers at Florida State University from an introductory philosophy course.

One group of participants received the following vignette:

Actor Condition:

A trolley is hurtling down the tracks. There are five workers on the track ahead of the trolley, and they will definitely be killed if the trolley continues going straight ahead since they won't have enough time to get out of harm's way. There is a spur of track leading off to the side where another person is working. The brakes of the trolley have failed and there is a switch which can be thrown to cause the trolley to go to the side track.

Imagine that you are an innocent bystander who happens to be standing next to the switch. You realize that if you do nothing, five people will definitely die. On the other hand, you realize that if you throw the switch, you will definitely save the five workers. However, you are also aware that in doing so the worker on the side track will definitely be killed as the result of your actions.⁴⁷

Participants in the actor condition were then asked the following questions:

1. Even if you don't think you have a moral obligation to throw the switch, is it nevertheless morally permissible for you to kill the one in order to save the five? Yes No
2. On a scale of 1 to 7, please state how much control you think you have over the outcome —1 being absolutely no control, 7 being complete control:

(no control) 1 2 3 4 5 6 7 (complete control)

The other group was in the Observer Condition. They received the same vignette except the "you" was replaced with "John." Upon reading this alternate scenario, participants were asked the following questions:

1. Even if you don't think that John has a moral obligation to throw the switch, is it nevertheless morally permissible for him to kill the one in order to save the five?
Yes No
2. On a scale of 1 to 7, please state how much control you think John has over the outcome

46 Attributional biases are tendencies to give different ratings of responsibility for a consequence of an action depending on whether one is the person performing the action or one is observing another person performing the action.

47 We purposely included both "kill" and "save" language to ensure that a word framing effect would not confound our results.

—1 being absolutely no control, 7 being complete control:

(no control) 1 2 3 4 5 6 7 (complete control)

Sixty-five percent of participants in the actor condition judged that it was morally permissible for them to hit the switch, but 90% of participants in the observer condition judged that it was morally permissible for John to hit the switch.⁴⁸ Similarly, whereas the average control rating for the participants in the actor condition was 4.3, the average control rating for participants in the observer condition was 5.1.⁴⁹ These results suggest that (a) people judge that it is less morally permissible for them to hit the switch than it is for others to hit the switch, and (b) people attribute less control over the outcome when they are the actors than they do when they are observers.

The important thing to note is that the only difference between the cases is the first and third person perspective, and the different perspectives alter intuitions about the scenarios. This suggests that whether people are actors or observers influences some people's intuitions. But these frames should not change intuitions about identical actions. That is contrary to the idea that moral intuitions track perspective independent moral facts. If people's intuitions only track perspective independent moral facts, then there should be no differences in responses. But there are. Hence, we have good reason to think that people's intuitions do not only track agent-independent moral facts. This is troubling for an objectivist who thinks that if all the morally relevant factual features are the same and only the actors are different, then the moral status of the action should not change.⁵⁰ Our results indicate that people's intuitions do not behave that way. If this is right, then we have even more reason to suspect moral intuitions are not reliable.

3. More Empirical Evidence about Moral Intuitions

3.1 Objectivity and Individual Differences

In an interesting research project, Donelson Forsyth (1980, 1981, 1992) has found that there are stable individual differences in how people respond to moral scenarios. He finds that people tend to vary along two dimensions—idealism and relativism. Highly idealistic people

48 These results are statistically significant: $\chi^2(1, N=85) = 7.873, p=.005$.

49 These results are statistically significant: Actor Condition $M=4.2791$, Observer Condition $M=5.1190, t(83)=-2.217, p=.029$.

50 One might argue that this also suggests that the supervenience platitude offered by Smith is also not a platitude held by the folk. In the actor-observer cases, all the factual features are the same and the only thing that changes is the perspective. If moral facts supervene on non-moral facts, then there should be no change in intuitions. But there is.

think desirable consequences can always be obtained, whereas people who are not idealistic think that bad consequences are often mixed with desirable results. Relativists reject universal moral principles whereas non-relativists think there are absolute moral principles. One could be high or low in idealism and high or low in relativism (Forsyth, 1980, p. 176). This gives us four basic ethical ideologies: (a) situationists who are high in relativism and idealism; (b) subjectivists who are high in relativism but low in idealism; (c) absolutists who are low in relativism yet high in idealism; (d) exceptionists who are low in relativism and low in idealism (Forsyth, 1992, p. 462). Research suggests that these four ideologies are associated with intuitions about “the ethics of caring (Forsyth, Nye, & Kelley 1988), business ethics (Forsyth, 1992), conformity to social norms and importance of consequences (Forsyth, 1985), and judgments of experimental ethics (Forsyth & Pope, 1984).⁵¹

In addition, Dollinger and LaMartina (1998) have found that those who are open to experience (one of the Big Five personality traits) reason differently about ethics. People who are open to experience tend to be more receptive to experience, less likely to reason in accordance with accepted societal standards, and are less likely to take for granted moral wisdom passed on by authority (Dollinger & LaMartina, 1998, p. 351). Because those who are open to experience are less likely to take external standards or societal norms as reasons for moral decisions, they predicted that those who are open to experience are more likely to use principled moral reasoning as opposed to conventional or socially accepted moral reasoning. Using Rest's (1979) Defining Issues Test, Dollinger and LaMartina found that those who are open to experience are more likely to engage in “post-conventional” moral reasoning (Kohlberg, 1969).⁵² Post-conventional moral reasoning is principled and is less reliant on threats, rewards, punishments, or duties to obey accepted moral standards of one's society.

Given the evidence that individual differences are important in moral intuitions and that openness to experience is correlated with some types of moral reasoning, Cokely and I (Feltz & Cokely, manuscript) hypothesized that those who are highly open to experience would be more likely to express non-objectivist intuitions than those who are low in openness to experience. Because those who are open to experience are more likely to reason individualistically, we

51 Interestingly, these ideologies are not associated with different behaviors, but absolutism is associated with harsher self-appraisals for moral norm violations (Forsyth & Berger, 1982).

52 The Defining Issues Test gives participants a series of moral dilemmas. They are then “asked to rank order and make ratings for various arguments, judging their ultimate importance in solving each moral dilemma” (Dollinger & LaMartina, 1998, p. 350).

thought it would be likely that they would be more skeptical that there is a definitive answer to moral disputes. They would be more open to the possibility that there is no single fact of the matter about moral reality. This led us to think that those who are open to experience would be less likely to be objectivists about ethics.

To test our hypothesis, participants were given personality and cognitive reflection measures along with both of Nichols' (2004b) moral and world scenarios, counterbalanced for order (See Appendix H). One hundred and twenty students at Florida State University from lower-level philosophy classes volunteered to participate in the experiment. Participants were allowed to respond in one of three ways. They could respond that the first person in the debate is right and the other is wrong, that the second person is right and the first is wrong, or they could respond that neither one is right because there is no fact of the matter. Those who responded that one of the two people in the debate is right were coded as objectivists, and those who responded that neither is right were coded as non-objectivists.⁵³

Five participants were excluded. Four participants requested their answers not to be used because they did not take the experiment seriously (they were too tired, did not read the questions, randomly selected answers), and one participant responded to the world scenario but not to the moral scenario. We found a substantial number of people ($N=79$, 69%) gave the non-objectivist answer to the moral scenario while a minority ($N=36$, 31%) gave an objectivist answer, a significant result.⁵⁴ Confirming previous research, we also found that significantly more people are objectivists about physical facts ($N=94$, 81%) than those who thought there is no fact of the matter about whether the earth is flat ($N=22$, 19%).⁵⁵

While the result that most undergraduates respond as non-objectivists about ethics is interesting (and replicates Nichols's findings), our primary concern was whether stable individual differences account for these responses. They do. As predicted, those who scored high in openness to experience were much more likely to respond as non-objectivists to both the moral and world scenarios than those who scored low. Splitting the groups into upper and lower

53 There are worries about categorizing participants this way. See Nichols 2004 for a compelling defense of this method.

54 $\chi^2(1, N=115) = 16.078, p < 0.01$. Even when we exclude those who answer as “global” non-objectivists as Nichols does, we still find significantly more participants giving non-objectivist answers to the moral question. Fifty-nine of 93 participants who gave an objectivist answer to World gave a non-objectivist answer to Moral, still a significant result $\chi^2(1, N=93) = 6.720, p < 0.01$.

55 $\chi^2(1, N=116) = 44.690, p < 0.01$

quartiles (extreme groups analysis), those who were high in openness to experience (N=28) were much more likely to give non-objectivist answers than those who were low in openness to experience (N=31). Of those high in openness to experience, 23 gave a non-objectivist answer whereas 5 gave an objectivist answer, a significant difference.⁵⁶ For those who were low in openness to experience, there was no significant difference between those who gave a non-objectivist answer (N=15) and those who gave an objectivist answer (N=16).⁵⁷

While these results may not generalize because the sample comes from an anecdotally non-objectivist population, it does lend prima facie support to the idea that there are a substantial number people who are non-objectivists about morality. Most of the people surveyed think there is no fact of the matter about some ethical claims. Moreover, it appears that there are stable groups of people who have objectivist or non-objectivist intuitions in these cases.

3.2 Framing Effects and Individual Differences

Because individual differences in personality play a role in moral intuitions, Cokely and I thought that other, non-personality based individual differences would also play a role in moral intuitions. Specifically, we thought that cognitive reflectivity may play a role in moral framing effects. Those who are higher in cognitive reflectivity are less likely to go with “gut” reactions and are more likely to take longer in reaching their final decision (Frederick, 2005). Hence, cognitive reflectivity can sometimes play a large role in decision making. For example, those who are highly cognitively reflective are more likely to choose a gamble with a higher expected value than a value they will certainly receive (Frederick, 2005, p. 40). Because those who are more cognitively reflective are more thoughtful, we thought they would be more influenced by actor-observer framing than those who are less cognitively reflective. We thought that those who are highly cognitively reflective may think longer and harder about why they should pull the lever. But in the observer condition, because they are not as motivated to think of reasons why others should act, they would be less likely to think that the third person is obligated to pull the lever.

In an attempt to test our hypothesis, we used a trolley style example (Feltz and Cokely, manuscript). Fifty-eight students in lower level philosophy courses at Florida State University volunteered to participate. Each participant was given one of the following scenarios⁵⁸:

⁵⁶ $\chi^2(1, N=28) = 11.571, p < 0.01$

⁵⁷ $\chi^2(1, N=31) = 0.032, p > 0.05$.

⁵⁸ These scenarios were created by Thomas Nadelhoffer. They are based on Bernard Williams's “Jim and the

First/Third Person Jim: You (Jim) find(s) yourself (himself) in the central square of a small South American town. Tied up against the wall are a row of twenty natives, most terrified, a few defiant, in front of them several armed men in uniform. A heavy man in a sweat stained khaki shirt turns out to be the captain in charge and, after a good deal of questioning which establishes that you (Jim) got there by accident while on a botanical expedition, explains that the natives are a random group of the inhabitants who, after recent acts of protest against the government, are just about to be killed to remind the other possible protesters of the advantages of not protesting. However, since you (Jim) are (is) an honored visitor from another land, the captain is happy to offer you (him) a guest's privilege of killing one of the natives yourself (himself). If you (Jim) accept(s), then as a special mark of the occasion, the other natives will be let off. Of course, if you (Jim) refuse(s), then there is no special occasion, and Pedro here will do what he was about to do when you (Jim) arrived, and kill them all. With some desperate recollection of schoolboy fiction, you (Jim) wonder(s) whether if you (he) got hold of the gun, you (he) could hold the captain, Pedro and the rest of the soldiers to threat, but it is quite clear from the circumstances that nothing of that kind is going to work: any attempt at that sort of thing will mean that you will also be killed along with all of the natives (including himself). The men against the wall, and the other villagers, understand the situation, and are obviously begging you (him) to accept. What should you (Jim) do?

And they were asked the following question:

Do you think that in these circumstances you (Jim) are (is) morally obligated to shoot and kill the one in order to save the others?

Yes No

We found no overall differences in responses between frames.⁵⁹ However, cognitive reflectivity interacted with the frames. We used the median cognitive reflectivity score to divide participants into groups that were high or low in cognitive reflectivity. Those who were more cognitively reflective tended to respond that one is more obligated to kill in the first person scenario than in the third person scenario. Table 9 represents the mean responses (0=no, 1=yes).

Indians" scenario (Smart & Williams, 1973).

⁵⁹ $F(1, 58) = .225, p > 0.05$

Table 9: Mean Responses to Jim and Natives

| | Low Cognitive Reflectivity | High Cognitive Reflectivity |
|--------------|----------------------------|-----------------------------|
| First Person | .41 | .92 |
| Third Person | .53 | .58 |

The interaction of frame with cognitively reflectivity was significant.⁶⁰ Indeed, cognitive reflectivity and frame explained a sizable amount of the variance.⁶¹ Hence, it looks like those who are highly cognitively reflective made judgments as though the difference in perspective changes moral obligations.

4. Implications of the Empirical Data

4.1 Critical Uses of the Empirical Data

The critic could argue that the evidence presented in this chapter points to serious problems with the reliability of moral intuitions. First, it looks like moral intuitions are not non-inferentially justified. The framing effects others have found and the framing effects reported for the trolley problem in this chapter indicate that intuitions can be influenced by morally irrelevant features. Given that all the morally relevant features of the cases remain the same, on these theories it should not matter if “I” perform the action or if another person does. According to some philosophers (e.g., Smith, 2004, Shafer-Landau, 2003), that shift in perspective should not change intuitions about what the right action is. But it does. If intuitions are illegitimately influenced by perspective, then they are not reliable. Hence, the evidence from these experiments indicate that intuitions, in and of themselves, are not reliable guides to moral truths.

Second, it looks like some philosophers are not good at determining, a priori, what moral intuitions the folk have. As the evidence suggests, some philosophers who think the folk are pre-theoretical objectivists about ethics appear to be wrong. There is a sizable number of people who

60 $F(3, 58) = 3.081, p < 0.05$

61 $\eta^2 = .146$

express non-objectivist intuitions about ethical cases. Furthermore, it does not appear that the folk endorse the practicality requirement on moral judgments. Given that philosophers sometimes are wrong about what moral intuitions the folk have, it is illegitimate for philosophers to make appeals to folk moral intuitions without the right kind of supporting empirical evidence.

Hence, it looks like folk moral intuitions are poor candidates as evidence for or constraints on our moral conceptual analyses and ethical theories. First, the evidence suggests that philosophers do a poor job at describing the folk intuitions. If intuitions are misdescribed, then they cannot legitimately provide evidence or constrain theories. In addition, if analyses and theories depend at least in part on us, then misdescribing our intuitions will result in mistaken theories and analyses. Second, it looks like intuitions are too unreliable even when they are properly described. The framing effects reported in this chapter call into question their reliability. This unreliability indicates that we have conflicting intuitions about morally identical scenarios, and we have no way of telling which intuitions are the right ones. If intuitions are unreliable in this way, they cannot provide evidence or constrain theories and analyses because we cannot non-arbitrarily decide which intuitions to use. Therefore, we cannot use moral intuitions as evidence or adequacy constraints for ethical theories and analyses of moral concepts.

4.2 Constructivist Uses of the Empirical Data

Are the critics right when they claim that intuitions are too unreliable to be used in philosophical theories and analyses? The constructivist can gladly accept Sinnott-Armstrong's argument that intuitions are only inferentially justified. Indeed, that is consistent with the idea that there are no general, a priori justifications for intuitions. Part of the constructivist's project is to provide the requisite empirical evidence on which we can base philosophical theory. Hence, while moral intuitions may not be fully self-justifying, they can be inferentially justified if the appropriate empirical evidence is provided.

It looks like there is exactly that kind of evidence to support constructivist projects. The evidence reported in this chapter suggests two things. First, it suggests that different philosophers may accurately describe different folk concepts or clusters of intuitions surrounding some moral concepts. For example, the data suggest that there are groups of people who express non-objectivist intuitions about ethics. Those people are characterized by having the stable personality trait openness to experience. While those who are low in openness to experience are more likely to judge that there are objective moral facts, those who are high in openness to experience are more open to the non-objectivity of morality. In some cases, those who are high

in openness to experience judge there is no single correct answer for moral disputes. If that is true, then that might indicate that those who are high in openness to experience have different, stable sets of moral intuitions from those who are low in openness to experience.

If there are different folk concepts, or at least stable differences in intuitions between different people, then it is true that philosophers are poor at determining *the* folk concept or folk intuitions because there simply is no single folk concept or set of intuitions. However, they may do a good job describing *some* ethical concepts or sets of intuitions. That is, there appear to be multiple folk concepts or sets of intuitions. When philosophers make reference to some of the folk concepts or sets of folk intuitions, it looks like what they are claiming is *true*. For example, for some people, it is a platitude that there are objective moral facts that determine the truth or falsity of ethical claims. Hence, Smith and Shafer-Landau are correct insofar as there are some groups of people who have the same intuitions they do. While this may not be what philosophers take themselves to be doing (they may indeed think they are offering an analysis of *the* folk concept), there are groups of folk who have the intuitions that support their analyses and theories.

Second, there is evidence that individual differences can help identify groups of people that are subject to framing effects. For example, cognitive reflectivity interacts with frames such that those who were highly cognitively reflective were more likely to say that one should kill the natives in the first person frame and less likely to say that Jim should kill the natives. However, low cognitive reflectivity did not interact with the frames. Hence, because individual difference can predict intuitions, we have good reason to think that there are different concepts or at least different, stable sets of intuitions about the cases. Hence, the apparent unreliability of intuitions is a surface phenomenon generated by subgroups of individuals who consistently express stable intuitions about the cases.

These data about framing effects and individual differences suggest that folk intuitions, while fragmented into different, identifiable groups, are stable enough to provide some evidential support for theories and analyses. One crucial premise in the critics' argument is that the unreliability of intuitions is a general feature of folk intuitions. But the evidence presented in this chapter indicates that folk moral intuitions are not generally unreliable in the sense that not all intuitions are influenced by irrelevant factors. Instead, folk moral intuitions are the result of stable features of groups of people. For example, people who are cognitively reflective are subject to the actor-observer bias in moral judgments. Indeed, if the results in this chapter are

right, then the actor-observer bias may not be a *bias*. Rather, the effect may be the result of an application of different concepts or reasoning strategies that lead some to think one is more obligated to throw the switch than one thinks others are.⁶² We cannot say that either of these reasoning processes is biased because we do not know the correct answer to these moral questions and because we do not have evidence that intuitions are generally unreliable. Absent those two pieces of information, it is premature to say that the processes that generate these intuitions are biased. In any event, these data provide the kind of inferential support that is required by Sinnott-Armstrong. Hence, while Sinnott-Armstrong may be correct that moral intuitions cannot be non-inferentially justified, we do have some inferential support for the evidentiary status of moral intuitions.

One might worry that if the arguments presented in this chapter are right, then that leads to an uncomfortable moral relativism. If our moral theories and conceptual analyses find evidential support in or are constrained by folk intuitions, then the fragmented nature of folk intuitions and concepts will lead to a fragmentation of ethical theory and conceptual analysis. And, because folk intuitions play these roles, it is possible that at the end of the day there will be nothing left to decide between some of these ethical theories or analyses. But that kind of relativism, the objection goes, is unacceptable in ethics.

This worry is legitimate, but its force can be mitigated. It is true that it is possible that at the end of the day we may have to adopt a kind of relativism. But that is true only if we want to use folk intuitions in one of the two ways mentioned at the beginning of this chapter. However, there is nothing necessitating that we use intuitions in those ways. Therefore, the correct move may be simply to reject those uses of folk intuitions.

While one may be tempted by this move, it does not come without costs of its own. If we reject the idea that folk intuitions are important or relevant to philosophical debates, then we run the risk of having our moral theory be nothing more than a “philosophical fiction” (Mele, 2001, p. 27). Considering that every culture on earth has a set of moral values (and some of them are shared), it would be a considerable cost to say that none of these people's intuitions are relevant to philosophical debates. Not only that, but it would be incumbent on those who reject folk intuitions to explain (a) why we should reject folk intuitions and (b) why people are mistaken

62 One might worry that we ask about “obligation” and the framing effect offered by Nadelhoffer and Feltz focused on “permissibility.” Because obligation entails permissibility, we thought that we would most likely find an effect on a stronger moral notion. Hence, we decided to focus on obligation and not permissibility. It is still an open question, however, if some individual differences account for the framing effect on permissibility.

about their moral views.

It is an open question whether the folk are really so fragmented. Another possible interpretation of the data presented here is that all people basically have the same moral intuitions or moral concepts. The different responses may be the result of different interpretations of the scenarios based on differences in individual goals, knowledge, or motivations. These differences are known to vary in predictable ways and in ways that are correlated with individual differences. For example, those who are high in openness to experience may have different goals or knowledge. They may be motivated not to reason in accordance with accepted standards, and the interplay of these features may result in systematically different responses. This leaves open the question what concept those who are open to experience have. Hence, it is possible that we will not have to be relativists or reject the importance of folk intuitions.

Hence, the constructivist agrees with the critic that there needs to be some inferential support if we are to use folk intuitions as evidence or constraints, and some research needs to be done to understand what we think about morality if analyses or theories are in part dependent on us. However, the constructivist thinks that when the requisite evidence is provided, we find that intuitions are not as unreliable as some critics suggest. There is widespread agreement in intuitions, and what appeared on the surface to be an unreliability in fact turns out to be a complex stability. If, as the constructivist suggests, intuitions end up being reliable, then we can use them for evidence and as constraints.

5. Conclusion

I have argued that on many points the critic is right: (1) there can be no non-inferential justification of moral intuitions, and (2) philosophers are not very good at determining, a priori, what moral intuitions or concepts the folk have. However, I argue that the critic overreaches when she claims (3) folk moral intuitions are not reliable enough to be used in ethical theories and moral conceptual analyses. The constructivist can agree with (1) and (2), and yet deny (3). The constructivist argues the empirical evidence suggests that moral intuitions are stable enough for conceptual analyses and ethical theorizing. That is, the data that call into question the reliability of folk intuitions reflect a surface phenomenon. When intuitions are more deeply probed, stability is revealed at group levels. This group stability is sufficient to underwrite using those stable intuitions as a basis for ethical theories and moral conceptual analyses. Hence, moral intuitions are reliable enough to serve as evidence for ethical theories and moral conceptual analyses.

CHAPTER 5: FREE WILL AND EXPERIMENTAL PHILOSOPHY

It seems if any field is insulated from the concerns of the experimentalist, it would be the study of free will. After all, compatibilists forward the metaphysical thesis that free will and moral responsibility are compatible with the truth of determinism, whereas incompatibilists hold that free will and moral responsibility are incompatible with the truth of determinism.⁶³ Because these are *metaphysical* theses, whether the folk have the intuition that we have free will is unrelated to whether we are in fact free and responsible. It is not strange to say that people can think they have free will and be wrong. Some philosophers expressly endorse such a position (Smilansky, 2002). Because folk intuitions are unrelated to whether we in fact have free will or not, it is not clear that folk intuitions can be used as evidence or adequacy constraints for theories of free will and moral responsibility.

The experimentalists gain a foothold in the fields of epistemology, ethics, and action theory because those fields are in part defined by what *we* think. For example, it would be odd if most people were systematically wrong in using the word 'intentionally' or 'knows'. However, it is not odd to say that people are systematically wrong when they say that one is free or responsible. Hence, at least on these grounds, it may seem that there is no pressure to try to keep analyses of free will in line with the everyday conception of free will.

Despite these reasons for thinking folk intuitions about free will and moral responsibility have no bearing on issues in free will, experimental results do have some role to play in the free will debate. First, experimentalists have impacted the debate as critics. A number of philosophers have made seemingly empirical claims about the folk position on free will without offering any empirical support for those claims. Second, experimentalists have attempted to support some positions with data about folk intuitions. For example, some philosophers have argued that philosophical views that are not in accord with folk intuitions shoulder the additional burden to explain why most people's intuitions are wrong. In this chapter, I offer some examples of philosophers who claim their position is supported by the folk view. Empirical evidence is offered that calls into question some of these claims. I also review some attempts to use folk intuitions about free will and moral responsibility to support philosophical arguments, and I raise

⁶³ Determinism is the thesis that 'at any instant exactly one future is compatible with the state of the universe at that instant and the laws of nature' (Mele, 2006, p. 3).

the worry that folk intuitions may not be reliable enough to be used in this role. I conclude that there is hope that intuitions about freedom and moral responsibility are stable enough to support the constructivists' project.

1. The Basic Positions, Intuitions, and Criticism

Traditionally, the positions on free will divide into two general camps—compatibilists and incompatibilists. Compatibilists think that free will and moral responsibility are compatible with the truth of determinism. Incompatibilists think that free will and moral responsibility are not compatible with the truth of determinism.

The first attempt at using the experimental method in the free will arena was done by Eddie Nahmias, Stephen Morris, Thomas Nadelhoffer, and Jason Turner (NMNT) (2004). They pointed out some philosophers use common, folk intuitions about free will and moral responsibility to support their philosophical views. If the view is in line with the folk view, then the view is supported because it is commonsensical in a way that a competitor's view is not. For example, Robert Kane writes that “most ordinary persons start out as natural incompatibilists” (1999, p. 217). Others, like Daniel Dennett, think that the folk are compatibilists because it is irrelevant “whether the agent in question could have done otherwise in the circumstances” (1984, p. 558). However, as NMNT argue, often philosophers have their intuitions contaminated by their theories (2004, p. 163).⁶⁴ Because what are used, in part, to support these philosophical views are folk intuitions, we should not look at philosophers who have spent years studying these issues. Rather, the relevant intuitions are the ones that are free of prolonged theoretical reflection. Hence, the data to be analyzed, at least in this respect, are folk intuitions and not philosophical intuitions.

NMNT argue that if these philosophers are to get support from folk intuitions, it had better be the case that the philosophers' intuitions match up with the folk intuitions. This is an empirical matter that cannot be verified from the armchair. Hence, an empirical investigation into folk intuitions is necessary for those philosophers who appeal to the common understanding of free will as supporting evidence. According to NMNT, once the folk conception of free will is understood, then that will “situate the burden of proof: if libertarian descriptions of our experiences are right, then compatibilists must explain why it shouldn't matter if those experiences are illusory, and if compatibilists' descriptions are right, then libertarians must

⁶⁴ The observations of NMNT suggest that some philosophers fall prey to the false consensus effect—the tendency of people to think that most others agree with them.

explain why we need to satisfy conditions for free will more demanding than what is suggested by our experiences” (Nahmias et al., 2004, p. 164).

Part of one prominent compatibilist strategy is the Conditional Analysis of “could have done otherwise,” that implies roughly that one performs action *A* freely and is morally responsible for performing it if the following is true: if one had decided to perform some action *B*, rather than *A*, one would have *B-ed*. That means “a different outcome—action *B*—would have ensued had the causal chain leading to action *A* been different” (Berofsky, 2002, p. 182). There are a host of philosophers who hold some variation of the conditional analysis of could have done otherwise, such as Hume, Hobbes, Schlick, Moore, and Ayer (Berofsky, 2002, p. 182-3). Indeed, some philosophers such as Adolf Grunbaum and J.S. Mill contend that the conditional analysis of could have done otherwise captures the folk conception of could have done otherwise (Nahmias et al., 2004). The conditional analysis of could have done otherwise, these compatibilists hold, is an important component for one being free and morally responsibility.⁶⁵ However, incompatibilists think that one is free only if one could have done otherwise even if nothing in the past was different.

NMNT explored intuitions about how people experience this ability to do otherwise. In these studies, NMNT gave the following scenario to 96 undergraduate participants:

Imagine you’ve made a tough decision between two alternatives. You’ve chosen one of them and you think to yourself, ‘I could have chosen otherwise’ (it may help if you can remember a particular example of such a decision you’ve recently made). Which of these statements best describes what you have in mind when you think, ‘I could have chosen otherwise’?

A. ‘I could have chosen to do otherwise even if everything at the moment of choice had been exactly the same’.

B. ‘I could have chosen to do otherwise only if something had been different (for instance, different considerations had come to mind as I deliberated or I had experienced different desires at the time)’.

C. Neither of the above describes what I mean. (2004, p. 174)

NMNT report that 62% answered 'B', a compatibilist friendly answer, while just 35% answered 'A', an incompatibilist friendly response (2004, p. 175). These results give *some* reason to think

⁶⁵ Of course, some additional sufficient conditions may be needed, such as *A-ing* is done in the right way (e.g. in the absence of constraints and manipulation). But this sketch is enough to illustrate at least how the compatibilist can see free will and moral responsibility are compatible with determinism.

that the folk do not operate with an incompatibilist notion of “could have done otherwise,” and hence *some* support that the more robust incompatibilist notions of could have done otherwise are not necessary for free will and moral responsibility.

NMNT performed more studies to get a clearer picture of what the folk intuitions about free will are. Instead of focusing on the phenomenology of free will, they conducted studies to see what intuitions the folk would have about various scenarios describing determinism. NMNT claimed that “if incompatibilism is not the intuitive view, or if no premises that support incompatibilist conclusions are particularly intuitive, then there seems to be little motivation for advancing an incompatibilist theory of free will” (Nahmias et al., 2006, p. 32-3). First, incompatibilism is a more metaphysically demanding view because it requires the falsity of determinism for freedom and moral responsibility—something compatibilism does not require. If it is more metaphysically demanding and not needed to explain common notions of free will and moral responsibility, then there seems to be little reason to endorse incompatibilism. Second, if incompatibilism is not the folk view, then incompatibilists will no longer be able to use folk intuitions to shift the burden of proof to the compatibilist (Nahmias et al., 2006, p. 29). Hence, “if it turns out that incompatibilist theories are not nearly as intuitive as incompatibilists themselves commonly assume, then it becomes increasingly difficult to see why we should adopt these theories” (Nahmias et al., 2006, p. 33).

To determine if incompatibilism is intuitive, NMNT test the following incompatibilist prediction:

(P) When presented with a deterministic scenario, most people will judge that agents in such a scenario do not act of their own free will and are not morally responsible for their actions. (Nahmias et al., 2006, p. 36)

(P) is tested against a variety of scenarios: (1) A Laplacean demon scenario; (2) a playing back the tape of history scenario; (3) a determinism by genes and upbringing story. In (1), NMNT tested a morally bad action (robbing a bank), a morally good action (saving a child), and a morally neutral action (going jogging). Seventy-six percent, 69%, and 79% of the participants, respectively, thought the person in (1) acted of their own free will. In response to the moral responsibility of the person (1), 88% responded that the person was morally responsible for the good action, and 83% thought the person was morally responsible for the bad action. In (2), participants were given a scenario where a woman stole a necklace. Sixty-six percent of the participants thought she did so of her own free will, and 77% thought she was morally

responsible. As for (3), they were given a scenario describing twins separated at birth and given different parents. One day, both twins find a lost wallet. One twin, Barney, returned the lost wallet, while the other twin, Fred, kept the wallet. Seventy-six percent of participants answered that Fred and Barney acted of their own free will. Sixty percent of participants thought Fred was morally responsible and 64% thought Barney was morally responsible. “The results from these three studies offer considerable evidence for the falsity of the incompatibilist prediction (P)” (Nahmias et al., 2006, p. 39). Hence, it looks like the incompatibilist shoulders the burden of proof because incompatibilism appears not to be the commonsense position.

Shaun Nichols and Joshua Knobe (in press) argued that compatibilist intuitions can be generated by what type of scenario is used. Specifically, they claim that “when people are confronted with a story about an agent who performs a morally bad behavior, this can trigger an immediate emotional response, and this emotional response can play a crucial role in their intuitions about whether the agent was morally responsible” (Nichols & Knobe, in press). To substantiate this claim, they ran two different sets of experiments. The first experiment tested folk intuitions about cases that are either abstract or concrete. Abstract scenarios are ones which are “designed to trigger abstract, theoretical cognition” and concrete scenarios are ones which are “designed to elicit greater affective response” (Nichols and Knobe, in press).

The participants were presented with the descriptions of the following two universes

Universe A: Imagine a universe (Universe A) in which everything that happens is completely caused by whatever happened before it. This is true from the very beginning of the universe, so what happened in the beginning of the universe caused what happened next, and so on right up until the present. For example, one day John decided to have French Fries at lunch. Like everything else, this decision was completely caused by what happened before it. So, if everything in this universe was exactly the same up until John made his decision, then it *had to happen* that John would decide to have French Fries.

Universe B: Now, imagine a universe (Universe B) in which *almost* everything that happens is completely caused by whatever happened before it. The one exception is human decision making. For example, one day Mary decided to have French Fries at lunch. Since a person's decision in this universe is not completely caused by what happened before it, even if everything in the universe was exactly the same up until Mary made her decision, it *did not have to happen* that Mary would decide to have French Fries. She could have decided to have something different. (Nichols and Knobe, in press)

Universe A is a description of a deterministic world, and Universe B is a description of an indeterministic world. The participants were then prompted to consider whether in Universe A a man who sets fire to his house and kills his family (concrete condition) is morally responsible.⁶⁶ Seventy-two percent of the participants said that this man was morally responsible (Nichols & Knobe, in press). However, when given the following abstract condition, “In Universe A, is it possible for a person to be fully morally responsible for their actions?” 86% percent of the participants said “no” (Nichols & Knobe, in press). Given this, it looks like in certain contexts people can have either compatibilist or incompatibilist intuitions.

Nichols and Knobe tentatively suggest the best explanation of the data is an “affective performance error.” The affective performance error explains the differences in judgments by positing a bias that is generated when participants are “faced with truly egregious violation of moral norms (our concrete cases), [where] they experience a strong affective reaction which makes them unable to apply the theory correctly” (Nichols & Knobe, in press). Nichols and Knobe claim that this is especially problematic when the people in the scenarios are described determinately (Nichols & Knobe, in press). They offer some empirical support for their interpretation. In a follow up study, they gave participants descriptions of Universe A and B. They also gave participants only concrete cases. The participants were given a person in either Universe A or B. Paired with one of these Universes was one of the following prompts:

High Affect: As he has done many times in the past, Bill stalks and rapes a stranger. Is it possible that Bill is fully morally responsible for raping the stranger?

Low Affect: As he has done many times in the past, Mark arranges to cheat on his taxes. Is it possible that Mark is fully morally responsible for cheating on his taxes? (Nichols & Knobe, in press)

In the indeterministic scenario, participants judged that it was possible that the person in the scenario was fully morally responsible (High 95%, Low 89%). However, in the deterministic universe, affect played a strong role. In the high affect case, 64% of the participants judged that it was possible that the person was fully morally responsible, but only 23% of participants did so in

⁶⁶ It is worth pointing out that even if Nichols and Knobe are right that affect can help generate compatibilist intuitions in cases with high affect, the affective performance error model does not explain why folk tend to give compatibilist responses to low affect cases. After all, according to the studies from the NMNT, 79% of people responded that going jogging can be done freely in a deterministic universe (2006, p. 39). Because jogging appears not to be affectively charged, the affective performance error model seems unable to explain that result. Of course, they may think that high affect is sufficient, in most cases for most people, to generate compatibilist intuitions but not necessary. But this provides some evidence that more than affect is doing most of the work in generating compatibilist intuitions.

the low affect case.⁶⁷ These data suggest that affect plays a role in biasing the subjects' intuitions about the person in the scenario. If the affective performance error model is correct, then the compatibilist responses from concrete high affect cases should be taken with great care because they are the result of a bias (Nichols & Knobe, in press).

2. Additional Evidence and Intuitions' Stability

It looks like the affective performance error model casts some doubt on NMNT's claim that most folk are compatibilists. However, there is still one interesting question about both NMNT and Nichols and Knobe's studies. If we look at the percentages of both NMNT and Nichols and Knobe's studies, we see that there is still a fairly robust dissenting minority. A natural question to ask is what explains this sizable minority? One explanation presents itself. It could be that *the* folk are neither compatibilists nor incompatibilists. Rather, it could be that some groups of folk are compatibilists and other groups of folk are incompatibilists.

To test this hypothesis, Edward Cokely, Thomas Nadelhoffer, and I (Feltz, Cokely, & Nadelhoffer, in press), ran Nichols and Knobe's experiment again. Instead of using a between subjects design, we used a within-subjects design. The prediction was that using a within subjects design would provide evidence that there are groups of people who answer the same way across questions about moral responsibility. In our first study, 52 undergraduates at Florida State University volunteered to participate in the survey. They were given both high and low affect scenarios, counterbalanced for order. Surprisingly, 25% of participants gave compatibilist answers to both questions, 67% gave incompatibilist answers to both questions, and 8% gave mixed answers. Significantly more participants provided incompatibilist answers to both questions.⁶⁸ In addition, we found no order effect depending on the presentation of the cases.⁶⁹ That is, judgments were stable regardless whether high affect came first or second. These data suggest that the folk are neither incompatibilists nor compatibilists, but both. If we can generalize from this sample, about 25% of the population are pre-theoretical compatibilists while about 67% of the population are pre-theoretical incompatibilists.

Of course, one major flaw with Nichols and Knobe's studies is that they only asked

67 One possible reason why there is a difference between Low Affect and NMNT's jogging scenario is that Nichols and Knobe use the "has to happen" language to describe determinism and NMNT do not. This possibility is discussed more fully in section 2.

68 $\chi^2(1, N=48) = 10.083, p < 0.01$

69 All F values were less than 1.

participants about moral responsibility and are therefore not allowed to make any conclusions about free will. At best, they are entitled to claim that most folk have intuitions that moral responsibility is not compatible with determinism. So, we ran another set of studies to probe folk intuitions concerning the relationship between free will and determinism. Participants received the description of universe A and both of the following questions which were counterbalanced for order:

FW High: Bill lives in Universe A. As he has done many times in the past, Bill stalks and rapes a stranger. Does Bill rape the stranger of his own free will?

FW Low: Mark lives in Universe A. As he has done many times in the past, Mark arranges to cheat on his taxes. Does Mark cheat on his taxes of his own free will?

We hypothesized that participants would not give significantly different responses to FW high and low than previous participants did to the moral responsibility high and low affect conditions.

In our first experiment, 65 undergraduate students from low level philosophy courses at Florida State University volunteered to participate. The results from our revised study closely track the results of our earlier one. When participants were asked specifically about free will, 29% ($N=19$) gave compatibilist matched responses, 62% ($N=40$) gave incompatibilist matched responses, and 9% ($N=6$) gave mixed responses. The difference between matched compatibilist and incompatibilist responses was statistically significant.⁷⁰ Moreover, there was no significant general order effect present,⁷¹ and there was not an order effect with respect to the matched answers.⁷² Finally, we tried to keep track of 'changed answers' by asking participants to put an 'X' through their original answers and to circle their new answer. This allowed us to loosely measure whether people were reconsidering their answers. In this study, only three participants changed their answers. Whereas two participants changed their answers in the FW low affect condition from 'yes' to 'no' when the high affect condition was presented *second*, the other participant changed the answer from 'yes' to 'no' when the high affect condition was presented *first*.

70 $\chi^2(1, N=59)= 7.475, p<0.01$

71 $\chi^2(1, N=59)= 0.857, p>0.05$

72 When FW High is presented first, 18 participants gave incompatibilist responses whereas 22 gave incompatibilist responses when FW High is presented second, $\chi^2(1, N=40)= 0.4, p>0.05$. When FW High is presented first, 11 gave compatibilist answers while 8 gave compatibilist answers when presented FW High is presented second, $\chi^2(1, N=19)= 0.474, p>0.05$.

We ran a second study to replicate our results. This time, 110 undergraduate students at Florida State University volunteered to take part in the experiment. They received the same materials that the participants received in the first study. The results of this second study were nearly identical to those of the first—namely, 29% ($N=32$) gave compatibilist matched responses, 63% ($N=69$) gave incompatibilist matched responses, and 8% ($N=9$) gave mixed answers. Once again, there was a significant difference between those who gave compatibilist and those who gave incompatibilist matched responses.⁷³ There was neither a general order effect⁷⁴ nor an order effect with respect to the matched answers.

These results put additional pressure on NMNT's contention that most people are pretheoretical compatibilists. First, if the data from Nichols and Knobe are correct, then what accounts for some compatibilist responses is an affective performance error caused by the high affect of the concrete scenarios. Second, it looks like in cases where participants are given both high and low affect cases, there is no affective performance error. After all, 67% of participants gave an incompatibilist response to both questions. Given this evidence, it looks as if giving both high and low affect scenarios to participants simply reveals stable compatibilist or incompatibilist views, or perhaps for some individuals giving them both scenarios causes a reduction in the impact of the affective performance error. Third, in this sample only a minority of participants gave the compatibilist responses to both high and low affect cases. If all this is right, then it looks like most folk are pre-theoretical incompatibilists, but not all of them.

Of course, one might worry that the participants are just trying to make their answers match. If the participants understand that both questions are about a deterministic world, then there would be additional pressure on the subjects to match their answers or risk being perceived as “inconsistent.” One may object that because participants matching answers can account for the data, we have no reason to think the evidence we provide suggests that people have consistent compatibilist or incompatibilist intuitions about these cases.

There are at least three responses to this objection. First, there are some people who give different answers to the high and low affect scenarios. About 8% of participants in all three studies we performed gave mixed answers. Because some people give mixed answers, participants simply matching responses does not explain all the data.

Second, there are two ways to think about matching answers. One way to think about

73 $\chi^2(1, N=101)= 13.554, p<0.01$.

74 $\chi^2(1, N=101)= 2.784, p>0.05$.

matching answers is that whatever answer participants give to the first scenario will influence the way they respond to the second scenario. Yet, if Nichols and Knobe's data are correct, there should be a clear order effect present in the experiments. There is no general order effect present in the experiments, but there might be a systematic difference in responses depending on whether the high affect case is presented first or second. We should expect that when the high affect case is presented first, there would be more compatibilist matched responses; and when the low affect case is presented first, we should find more incompatibilist matched responses. When we look at the data, we find that this prediction is not supported. When the high affect condition was presented first ($N=26$), 65% gave incompatibilist responses, 31% gave compatibilist responses, and 4% gave mixed answers. When the low affect case was presented first ($N=26$), 69% gave the incompatibilist response, 19% gave the compatibilist response, and 12% gave mixed answers. Indeed, there is no significant difference between incompatibilist matched answers based on order.⁷⁵ There is also no significant difference with compatibilist matched responses based on order.⁷⁶ Given that there is no significant difference between those who gave incompatibilist and compatibilist answers based on the different orders of the scenarios, we have reason to think that people are not trying to match their second answer to their first. Likewise, we do not find a matched answer order effect in our studies about moral responsibility.⁷⁷

Another way to think about matching answers is that after the participants read the second scenario, they go back and revise their response to the first scenario. This possibility was foreseen. In the instructions to the surveys, the participants were told that if they decided to change an answer to one of the questions, then they should put an 'X' through their original answer and circle their new answer. On the moral responsibility survey, seven participants decided to change their answers (13%). Similarly, only three participants changed their answers in the first free will study, and only seven participants changed their answers in the second free will study. Hence, a small but non-significant number of subjects decided to change their answers.

Third, it could be that the subjects respond to the prompts after reading *both* scenarios.

⁷⁵ $\chi^2(1, N=34)=0.118, p>0.05$

⁷⁶ $\chi^2(1, N=14)=0.286, p>0.05$, although there is a small numerical difference in the predicted direction.

⁷⁷ Thirty-three gave incompatibilist answers when FW High was presented first, and 36 gave incompatibilist answers when FW High was presented second $\chi^2(1, N=67)=0.130, p>0.05$. 21 gave compatibilist answers when FW High was presented first, 11 gave compatibilist answers when FW High was presented second $\chi^2(1, N=33)=3.667, p>0.05$. It should be noted that there was a near significant effect for compatibilist matched answers.

The responses to the first two objections assume that the participants responded to the scenarios serially—an assumption which may be false. It is possible that the participants only answer the scenarios after reading both. This is something difficult to tease out in the current data set. However, it is important to note that the results of our experiments roughly mirror the results Nichols and Knobe found in their abstract condition. In their abstract condition, 86% of participants said that a person in a deterministic universe is not morally responsible. In the current studies, the results were not that robust. Sixty-seven percent of the participants reported that the people described in the scenarios were not responsible. In the free will studies, 62% and 63% said that the person was not free. Hence, a strong majority in both studies thinks that the people in a deterministic universe are not free or morally responsible. This correspondence between the studies suggests that these people, when given an opportunity to reflect on the cases, come to the conclusion that people in a deterministic universe really are not free regardless of the affect present in one of the cases.

Nichols and Knobe consider this possibility but find no evidence for it.⁷⁸ However, our study suggests that when participants are given an opportunity to review the high and low affect cases, they do not fall prey to the affective performance error. If the participants in our study are less likely to fall prey to the affective performance error, then we have reason to think that the majority of folk report intuitions as a result of theoretical reasoning and those intuitions are incompatibilist. It might be that the participants are not blindly matching their responses just out of some motivation to maintain the perception of consistency. As our data suggest that the highest percentage of participants is incompatibilist, the chance to reflect on two different cases may allow them to come to the reasoned position that incompatibilism is the correct position. Thus, even if the subjects are prone to an affective performance error when they are given high affect cases in isolation, it stands to reason that if participants give the same answers to both high and low affect cases, then it reflects their reasoned, non-performance error judgment.⁷⁹ If that is

78 They attempted to get people to adjudicate between the different intuitions in the following manner. They performed another experiment where the participants received information explaining the results of the previous studies. They are then told to attempt to say what the 'right' position is. Roughly half of the participants said that the compatibilist intuitions are correct and half said that the incompatibilist intuitions are correct (Nichols & Knobe, forthcoming).

79 While it is possible that some participants reflect on both questions before answering them, our evidence does not support that claim. There are roughly the same number of participants who change their answers in the different orders of the cases—five changed their answers when High Affect is presented first and two change their answers when Low Affect is presented first which is non-significant $\chi^2(1, N=7) = 1.286, p > 0.05$. Also, when looking only at the responses to the first question, our results do not replicate Nichols and Knobe's results as one would expect if affect is playing a role in judgments. That is, if people are prone to the affective performance

right, then we have even more reason to think that these responses are their true intuitions about free will and moral responsibility that are free from an affective performance error, and those intuitions do not disconfirm (P).

In relation to the previous worry, one might also worry that if the participants are thinking about these cases and coming to a reasoned conclusion, then we are not targeting their *pre-theoretical* intuitions. Rather, we are gathering their *theoretical* judgments about the cases. Hence, our data are beside the point because they do not tell us what we are really interested in—namely, whether the folk are pre-theoretical incompatibilists or compatibilists. While this objection is true as far as it goes, it is unclear how much traction it really has. First, it is not clear that the intuitions that NMNT and Nichols and Knobe gather are also not guilty. After all, Nichols and Knobe explicitly admit that the answers to the abstract condition are a result of theoretical reasoning (Nichols & Knobe, in press). Second, if we are interested in what the non-professional philosophers think about cases, then the intuitions gathered in our experiment are just as legitimate as the intuitions that are gathered by NMNT and Nichols and Knobe. Third, it is unclear why we shouldn't be just as interested in folk theories that generate these intuitions. Hence, whether the intuitions gathered in our study are theoretical is not a worry.

Of course, the results of these studies are not definitive. Eddy Nahmias (2006) has identified at least one additional problem with Nichols and Knobe's experimental design. Nichols and Knobe use a notion of “complete causation.” That is, in their description of Universe A, they say that everything in Universe A “has to happen” as it does. However, if things happen necessarily, then even the compatibilist would think we are not free and morally responsible (Nahmias, 2006, p. 223). The language of “had to happen” could be interpreted to mean that nothing the agent could have decided would have made any difference to her *A-ing*. But, on the conditional analysis of could have done otherwise, what an agent decides *could* make a difference to what action the person performs. Absent that ability to do otherwise, many compatibilists would say that the person in Nichols and Knobe's scenarios is not free. Because Nichols and Knobe's design uses a notion of determinism that even compatibilists think rules out

error, then we should expect that when High Affect is presented first a significant number of people would give a compatibilist response or would change their answer to the High Affect question after reading the Low Affect question. Neither one of these are found in our data. It is an interesting question what accounts for the differences in results between Nichols and Knobe's study and ours. One interesting thing to note is that 6 of the 7 people who changed their answers changed them from a compatibilist to incompatibilist responses—a nearly significant result, $\chi^2(1, N=7)=3.571, p=0.059$. This lends *some* credibility to the claim that some people, when given the opportunity to reflect on the questions, do indeed opt for incompatibilist over compatibilist responses.

free will and moral responsibility, we cannot conclude that the intuitions generated in these studies indicate anything at all about how the folk think freedom and responsibility are related to the relevant notion of determinism. Rather, these results may indicate how the folk think freedom and responsibility relate to fatalism—the thesis that things happen necessarily as they do (Feltz et al., in press.).

It appears this objection is correct. There is some evidence that people are in fact sensitive to the wording of the scenarios that imply fatalism. Nahmias, Coates, and Kvaran (NCK) (2007) conducted studies that described determinism in terms of complete causation yet did not use the “has to happen” language. These scenarios “described deterministic worlds in which agents' decisions are (1) completely caused by prior events, and (2) those prior events were completely caused by earlier events going back to events before the agent was born, such that (3) the prior events will definitely cause the later events” (Nahmias et al., 2007, p. 222). NCK tested to see if affect played a role in intuitions about concrete cases with the revised description of determinism that lacked the “has to happen” language. Affect did play some role. Agents who did something good were judged as less free and responsible than an agent who did something bad. But regardless of whether the action being performed was good or bad, about 60% participants responded that the person in the scenario has free will, and between 63-81% of participants thought the person was morally responsible (Nahmias et al., 2007, p. 227).⁸⁰ This suggests that simply leaving out the phrase that one's action “has to happen” leads to people having predominately compatibilist friendly intuitions.

In any event, for the present purposes, the point is that the evidence suggests there are at least two groups who express stable intuitions. If Nahmias's objection is correct, then one group expresses stable intuitions that fatalism is not compatible with freedom and responsibility, whereas another, much smaller group expresses stable intuitions that we are free even in a fatalistic world. Given that there is some stability of intuitions, it is an interesting question if there are individual differences which predict responses. We have some reason to think that we can predict responses on the basis of individual differences. Providing evidence for this additional stability is the task of the next section.

⁸⁰ NCK might have gotten different results from what Cokely, Nadelhoffer, and I got because people could be sensitive to the “has to happen” language that was present in our experiment but not in NCK's.

3. Framing Effects, Reductionism, and Individual Differences

There is some evidence that the way in which determinism is described influences the way in which people respond. In an interesting set of studies, NCK (2007) gave college undergraduates one of two structurally identical descriptions of a situation where determinism is true. One scenario is framed in psychologically reductionist terms where the mechanical processes of the brain are determined. The other scenario is framed in psychological non-reductionist terms where determinism is described in psychological terms. NCK used the following two versions—the non-reductionist version is in brackets, the reductionist is underlined:

Most respected neuroscientists [psychologists] are convinced that eventually we will figure out exactly how all of our decisions and actions are entirely caused. For instance, they think that whenever we are trying to decide what to do, the decision we end up making is completely caused by the specific chemical reactions and neural processes [thoughts, desires, and plans] occurring in our brains. The neuroscientists [psychologists] are also convinced that these chemical reactions and neural processes [thoughts, desires, and plans] are completely caused by our current situation and the earlier events in our lives, and that these earlier events were also completely caused by even earlier events, eventually going all the way back to events that occurred before we were born.

So, once specific earlier events have occurred in a person's life, these events will definitely cause specific later events to occur. For example, one day a person named John decides to kill his wife so that he can marry his lover, and he does it. Once the specific chemical reactions and neural processes [thoughts, desires, and plans] occur in John's brain, they will definitely cause his decision to kill his wife. (Nahmias, Coates, & Kvaran, 2007, p. 224)

Most people (about 60%) given the reductionist scenario thought the person was not free or responsible, whereas most people (about 85%) given the non-reductionist version did think the person was free and responsible (Nahmias, Coates, & Kvaran, 2006, p. 229).⁸¹ According to them, these results “suggest that as long as people are not primed to think that determinism entails mechanism, most do not perceive it as incompatible with FW [free will] or MR [moral

81 These results come from questions where “the agents were described abstractly” and not performing a particular, concrete action (Nahmias, Coates, & Kvaran, 2006, p. 225). NCK ran two version of these scenarios. One version described actions taking place on another planet Erta, and one describe actions taking place on earth. NCK only tested abstract questions for scenarios that took place on earth and they did not test concrete questions.

responsibility]” (Nahmias, Coates, & Kvaran, 2006, p. 232).

Cokely and I extended NCK's studies by testing concrete questions that took place on earth, and we hypothesized that stable individual differences could predict different groups of people who have the intuitions they do about NCK's scenarios. After all, there is a substantial dissenting minority in most of the studies conducted thus far. As well, Feltz et al. found that there are stable individual differences at the level of specific intuitions about free will and responsibility. So, it was likely that individual differences could in part explain the results of NCK's experiments. We examined these intuitions using Nahmias, Kvaran, and Coates's (2007) psychologically non-reductionistic scenarios for three reasons. First, those scenarios involve a socially important action of a man killing his wife. Second, they describe determinism in terms of complete causation while at the same time avoiding using terminology implying that events had to happen as they did.⁸² Third, we know that in the psychologically non-reductionist scenario the majority of people are compatibilists.

Answers of 1-3 were coded as agreeing with the statement, answers of 5-7 were coded as disagreeing with the statement, and an answer of 4 was coded as neutral. Agreement with prompts 1-3 indicates compatibilist intuitions, and disagreement indicates incompatibilist intuitions. Table 10 represents participants' responses:

Table 10: Responses to the Psychological Scenario

| | Compatibilists | Incompatibilists | Neutral |
|-------------|----------------|------------------|---------|
| Up to him | 71% | 26% | 3% |
| Free will | 69% | 19% | 2% |
| Responsible | 76% | 22% | 2% |

Replicating NCK's experiment, significantly more people expressed compatibilist intuitions.⁸³

⁸² See Feltz, Cokely, & Nadelhoffer, in press, or Turner & Nahmias, 2006 for a more detailed discussion.

⁸³ For the present analysis, we formed two groups. One group consisted of people who expressed compatibilist intuitions (responded 1-3) and the other of those who did not express compatibilist intuitions (those who responded 4-7). Significantly more people responded as compatibilists: Up to him $\chi^2(1, N=58) = 8.345, p < 0.01$; Free will $\chi^2(1, N=58) = 5.586, p < 0.05$; Responsible $\chi^2(1, N=58) = 22.345, p < 0.01$.

Three planned linear regression models were constructed using extraversion as the independent variable and agreement with the three different statements as the dependent variables. Consistent with our hypothesis, all models revealed significant relationships between extraversion and compatibilist responses. To further illustrate these relationships, planned analyses next followed a common approach in individual differences research and divided extraversion scores into upper and lower quartiles (extreme groups analysis)[Cokely, Kelley, & Gilchrist, 2006]. The means of these two groups are illustrated in Figure 1.

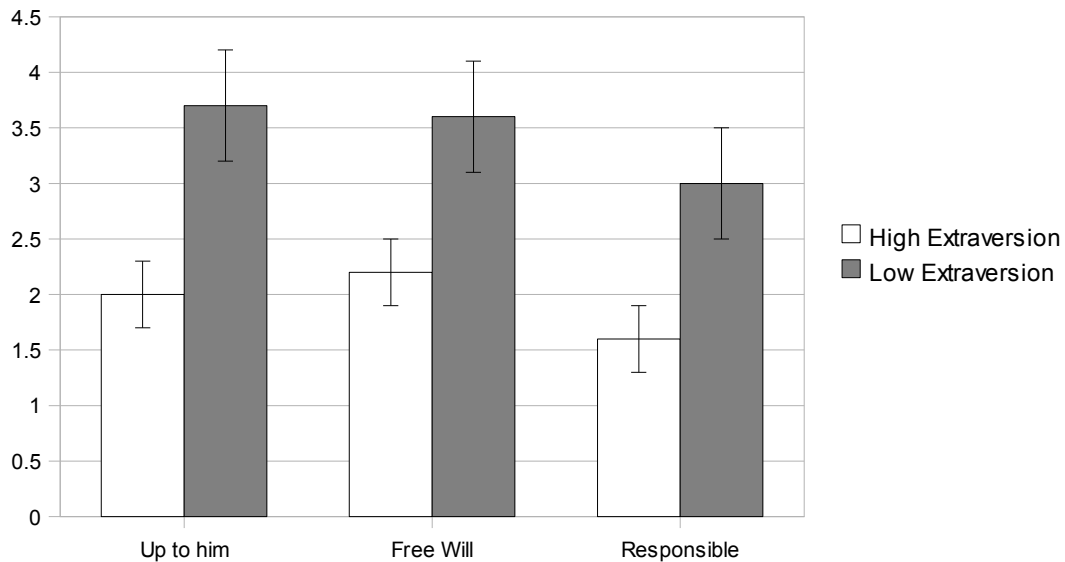


Figure 1: Mean scores by high and low extraversion.

As expected, the differences in responses between those high ($N=21$) and low ($N=20$) in extraversion are large and significant for all three questions.

These results demonstrate that extraversion is a reliable predictor of compatibilist and incompatibilist intuitions. First, the linear regressions indicate that extraversion is positively, linearly related to compatibilist judgments. Indeed, extraversion explains a sizable amount of the variance.⁸⁴ Second, when we look at those who are moderately high versus moderately low in

⁸⁴ Up to him: $\beta = -.38$, $t(56) = -3.073$, $p = .01$, $R^2 = .144$, $F(1, 56) = 9.441$, $p = .01$. Free will: $\beta = -.206$, $t(56) =$

extraversion, we find large qualitative differences in both (a) the up to him statement and (b) the free will statement.⁸⁵ Critically, those who are lower in extraversion tend to be neutral about (a) and (b) whereas those who are higher in extraversion agree to both. As well, there is a large quantitative shift between those who are high and low in extraversion in relation to the person being responsible. Those who are highly extraverted strongly agree that the person is responsible, but those who are low in extraversion only weakly agree.

4. Implications for the Free Will Debate

4.1 Critical Uses of the Evidence

A large part of the free will debate is about the relation of determinism to free will and moral responsibility. Because determining the nature of this relationship is a metaphysical thesis that may not be reflected in folk intuitions, some may think that folk intuitions do not have much purchase in the free will debate. So, one might think that because the nature of the free will debate is different from the debates in epistemology, ethics, and action theory, it is not obvious that the critic can make a direct argument from folk intuitions to philosophical intuitions. However, the critic could offer the following dilemma. Either philosophers (a) should pay attention to folk intuitions about freedom and responsibility or (b) should not pay attention to folk intuitions about freedom and responsibility. If (a) is right, then it is a short step for the critic to argue that the empirical results are troubling. First, let's look at the cases that describe determinism with "had to happen" language (Feltz et al., in press; Nichols & Knobe, in press). The "had to happen" language implies fatalism where things happen necessarily as they do. Because compatibilists would say that in fatalistic worlds people are not free and responsible, we cannot be sure that participants are responding as "true" incompatibilists. Participants may simply think that free will and moral responsibility are not compatible with fatalism, and those judgments say nothing about determinism's relation with free will and moral responsibility. However, a sizable minority still thought the people in the fatalistic scenarios were free and morally responsible, indicating that they think people are free no matter what. These people are normally discarded because they fail manipulation checks.⁸⁶ However, we cannot be sure that

$-2.125, p < .05, R^2 = .08, F(1, 56) = 4.515, p < .05$. Responsible: $\beta = -.201, t(56) = -2.0, p = .05, R^2 = .08, F(1, 56) = 4.0, p = .05$.

85 Up to him: $t(39) = 2.86, p < 0.01, d = .9$. Free will: $t(39) = 2.46, p < 0.05, d = .8$. Responsible: $t(39) = 2.27, p < 0.05, d = .6$.

86 Participants were also asked to answer 'yes' or 'no' to the following control question: "If the psychologists are right, is it accurate to say that if the universe were re-created, John would make the same decision"? An answer

some people get the manipulation checks right and still think we are free no matter what. So, we need more evidence that the “free will no matter what” people are not biasing results toward compatibilism. Without that evidence, we can't be sure what intuitions the folk have about the relationship of determinism to freedom and moral responsibility.

Apart from the worry that folk intuitions collected thus far may not reveal anything about determinism's relation to freedom and moral responsibility, there is a framing effect depending on how determinism is described. While NCK interpret their studies as suggesting people have the intuition that only some kinds of determinism are threatening to freedom and responsibility, it seems more likely that people's intuitions can be swayed by factors that should not be relevant to their intuitions about determinism's relation to freedom and responsibility. If everything in the world operates deterministically, then at what level determinism is described should not matter. That is, if we don't assume that people are composed of a body and a soul (in which case, the soul may always be free), it should not matter if physical states of our brains are determined or if the beliefs and desires that depend on those brain states are determined. But the results from NCK show that describing determinism differently does have an effect. Most people think that when determinism is couched in psychological terms, we can be free and responsible. However, when determinism is described in terms of physical states, most people think we are not free and responsible.⁸⁷ Clearly, it is not the determinism that is doing the work here, but something else. If there is this framing effect, we cannot be sure which intuitions accurately reflect people's “true” intuitions about free will and moral responsibility's relation to determinism.

Additionally, the critic could argue that if the results of Nichols and Knobe are right, then the intuitions that are generated in cases with determinately described agents may unduly bias participants to have compatibilist intuitions because of affect. If affect does not allow proper use of people's concept of free will and responsibility, then we should not have much confidence that these intuitions tell us something about their underlying concept of free will and moral responsibility. After all, if we can manipulate intuitions in this way, we just don't have a good way of determining what the “right” intuitions are. Are the intuitions about high affect cases the right ones, or are the ones about low affect or abstract conditions the right ones? It seems that we

of 'no' fails the control because if the universe were re-created, John would make the same decision.

87 Perhaps the two different ways of describing determinism activate two different concepts of causation which influence judgments. Many people may think that basic physical causation is not compatible with freedom and moral responsibility whereas more people think that psychological causation is. Participants may conceive of causation differently in these two different connections. I thank Al Mele for pointing out this possibility.

do not have a principled way to decide between them even if we say that the intuitions in the high affect cases are due to an affective performance error. Why aren't the intuitions in the concrete cases the result of a low affective performance error? In both cases the intuitions may result from an error. If we can't decide between them, then we should have little confidence that they are stable enough to inform us of concepts or serve as a basis for philosophical analysis.

If people are subject to a framing effect, then we have good reason to think that their intuitions are not reliable. If the results of Chapter 1 are right and there is no adequate a priori defense of intuitions, it would be question begging to assume that philosophers' intuitions are in any better position. Therefore, the critics claim, we have no reason to think that intuitions are reliable enough to be used as evidence for theories about free will and moral responsibility.

If (b) is true, then it is incumbent on the philosopher to explain why we might require a conception of freedom and responsibility that is above and beyond what most people think is required. As NMNT rightly point out, it would seem weird to demand a notion of freedom and responsibility that is more robust than the one called for by commonsense. If NMNT are right about folk intuitions, then all that we should require from our theories of free will and responsibility is a compatibilist notion and not the more metaphysically robust incompatibilist notion.

However, I have presented evidence in this chapter that indicates that there are groups of people who express stable intuitions about determinism's relation with free will and moral responsibility. If there are groups of people with stable intuitions, then *both* the compatibilist and the incompatibilist will have to take into account the empirical literature about folk intuitions. Both camps must take into account the literature because it looks like there are sizable minorities who express opposing intuitions about free will and moral responsibility. It is incumbent on any philosophical view about free will that attempts to garner evidential support from folk intuitions, then, to explain why those groups of people have the mistaken intuitions they do. The critic is at least in part satisfied by the fact that philosophers can no longer make appeals to folk intuitions from the armchair.

4.2 Constructivist Uses of the Empirical Results

In the previous section, I argued there are two experimental results which call into question the reliability of intuitions about free will and moral responsibility. The first result is that intuitions are influenced by affect. The second is that there is a framing effect on intuitions concerning free will and moral responsibility. Because of these results, I argue that the critic can

claim that folk intuitions about freedom and responsibility are not reliable. I also argue that we cannot assume that philosophers' intuitions are any more stable unless we beg the question. Because I agree with the critics' contention that philosophers can no longer just sit in their armchairs, in the present section I offer some ways that constructivists could answer these unreliability worries.

One way to counter the critics' arguments is to show that the unreliability of intuitions is only apparent. For example, what on the surface may look like a framing effect is really the result of an underlying stability. Different groups of people have stable, different intuitions, and those groups of people are identifiable by general individual differences. For example, Feltz et al. found that most people answered the same way to both the high and low affect conditions. One explanation of this is that they saw both questions before answering. Because of this chance to reflect, most people came to the reasoned position that in a deterministic universe, people are generally not free or morally responsible. If this is right, then that suggests that there is stability in intuitions when people are allowed to reflect on the cases. So, while intuitions can be manipulated, there are ways to reduce the effects of that manipulation.

In the replication studies of NCK, Cokely and I found that those who report themselves to be extraverted differentially answer questions in the non-reductionist frames. That is, they are less likely to think that the person in these studies is free. This indicates that the framing effect found in the two different cases might be the result of stable individual differences between subgroups of people. These results suggest that different groups have different, stable intuitions about these cases. Indeed, these results offer a refinement of NCK's view—it could be that only *some* people find determinism framed in non-reductionist terms not threatening to free will and moral responsibility. If there is this kind of stability, then the intuitions these people have are reliable.

If there are different groups with stable intuitions about these cases, that indicates two possibilities. First, it could indicate that these different groups have different concepts. If intuitions are in part driven by people's concepts, then differences in concepts would entail that intuitions about scenarios that engage those concepts would be different. Second, it could be that people interpret the cases differently. It is known that different people with different individual differences process, code, and retrieve data differently. This might also be true in the free will debate. People may simply code and retrieve data about scenarios concerning free will and moral responsibility differently depending on some individual differences. If this is true, then these

differences could account for the variability and apparent irrationalities in intuitions without necessarily having a different concept.

In any case, if the arguments in support of the constructivists are right, then we can offer a constructive analysis of folk concepts of free will and moral responsibility. And if folk intuitions are more reliable than critics might think, then the constructivist can block the inference that philosophical intuitions are problematic in the same ways as the folk intuitions. It could be that different individual differences predict what intuitions philosophers have. In turn, different philosophers could just have different concepts they are attempting to analyze. If this is true, then the constructivists have hope for an empirically informed analysis of determinism's relation to free will and moral responsibility.

5. Conclusion

In this chapter, I have reviewed a number of studies about intuitions in the free will debate. Some of these results support critical attacks on intuitions by calling into question the reliability of intuitions. However, using other results, I argue that the constructivist can counter the critical attack on intuitions in the free will debate. There are explanations for the framing effect, and there are a number of people who express stable intuitions across scenarios. If there is this stability, then we have good reason to think that the apparent unreliability of intuitions is really the result of differences in responses of sub-groups. These sub-groups can be identified and they appear to have stable intuitions. I conclude that if they have stable intuitions, then that could indicate that they have different concepts or that they interpret scenarios differently. Either one is sufficient to deflect part of the critical argument against intuitions in philosophy.

CHAPTER 6: INTUITIONS AND ACTION THEORY

Another area where the reliability of intuitions has been called into question is in action theory. According to Alfred Mele, “Central to the philosophy of action is a concern to understand *intentional* action” (1992, p. 199). The contribution of experimental philosophy to traditional action theory is important because it can help shed light on what the folk concept of intentional action is. As Mele correctly points out, “a philosophical analysis of intentional action that is wholly unconstrained by that [folk] concept runs the risk of having nothing more than a philosophical fiction as its subject matter” (2001, p. 27). Hence, without the requisite empirical data, we just cannot be sure that philosophical analyses are constrained by the folk concept of intentional action.

A number of philosophers have claimed that their analyses of intentional action are in accordance with the folk concept. For example, the Simple View (SV) states that intentionally doing action *A* entails intending to *A*. Hugh McCann thinks the SV “pertains to the everyday concept of intending” (McCann, 1998, p. 210). Fred Adams claims that the SV has a number of advantages which are embedded in the folk concept of intentional action: (1) it is a “simple” view that is not “cluttered” by unnecessary theoretical baggage; (2) it allows for intentions to play a direct causal role in production and sustaining of the action; (3) it allows intentions to be a guiding factor in an action's development and commission; and (4) it allows us to distinguish intentional from unintentional actions (1986, p. 284-5). He thinks, “these folk-action-theoretical principles, if you will, are deeply entrenched” (Adams, 1986, p. 285). Because the SV and the principles entailed by it are deeply entrenched and comport with the everyday concept of intentional action, there is strong *prima facie* reason to accept the SV.

However, it is an empirical question if the ordinary concept of intentional action includes anything like the condition indicated by the SV. There is reason to think it does not. Joshua Knobe (2003a) has discovered that the perceived goodness or badness of side effects influences people's ascriptions of intentionality. In this chapter, I use the Knobe effect as a case study of the ways in which folk intuitions relate to and inform analyses of intentional action. I first describe the Knobe effect, and I then explain some ways the critics and constructivists have used the Knobe effect. I present new evidence which supports and clarifies both the critics' and

constructivists' views. I conclude that intuitions involving the Knobe effect are for the most part reliable, and I offer ways these intuitions may point toward folk concepts of intentional action. If my arguments are right, and if Mele is right that analyses of intentional action should be in part constrained by folk concepts of intentional action, then folk intuitions play an indispensable role in the philosophical analysis of intentional action.

1. The Knobe Effect

If a consequence of an action is foreseen but not intended, then that consequence is a side effect.⁸⁸ The “Knobe effect” describes the tendency of people to judge a bad side effect is brought about intentionally whereas a good side effect is judged not to be brought about intentionally. The best known cases used to demonstrate the Knobe effect are Knobe's two chairman cases:

Harm

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

Help

The vice-president of the company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, and it will also help the environment.' The chairman of the board answered, 'I don't care at all about helping the environment. I just want to make as much profit as I can. Let's start the new program.' They started the new program. Sure enough, the environment was helped. (Knobe, 2003a, p. 191)

About 82% of the participants given Harm say that the chairman brought about the bad side effect (the harm to the environment) intentionally, while 77% of those given Help said the chairman did *not* bring about the good side effect (the help to the environment) intentionally (Knobe, 2003a, p. 192). These results have been replicated across a variety of cases involving side effects (Cushman & Mele, in press; Knobe 2003a, 2003b, 2004a, 2004b, 2006; Knobe &

⁸⁸ Mele and Cushman offer a more complete definition of a side effect: “X is a side-effect action performed by an agent S if and only if S successfully seeks to perform an action A, E is an effect of his so doing, X is his bringing about E, and X has the following properties: S is not at the relevant time seeking to X either as an end or as a means to an end, and X is not in fact a means to an end that S is seeking at the relevant time” (2007, p. 185).

Mendlow, 2004; McCann, 2005; Mele & Cushman, 2007; Nadelhoffer, 2004a, 2004c 2005, 2006a, 2006b, 2006c), cultures (Knobe & Burra, 2006), and ages (Leslie, Knobe, & Cohen, 2006).

2. Constructivist Explanations of the Knobe Effect

2.1 Single Core Concept View

Joshua Knobe (2006) attributes the asymmetric results to the core folk concept of intentional action. Knobe's (2006) model of intentional action ascription features two sub-processes. First, one determines if the behavior is good or bad. Second, one's concept of intentional action uses the determination from the first step to issue an intentionality judgment. To illustrate, consider judgments made about Harm. One first determines the side effect to be bad. Then, one's concept is engaged, and one searches for features of the bad side effect sufficient to judge it being brought about intentionally. For example, foresight might be sufficient to judge a harmful side effect is brought about intentionally. Hence, people think the chairman intentionally harmed the environment. However, a different set of considerations might be relevant when the chairman helps the environment—foresight may no longer be sufficient. What may be required in addition to, or instead of, foresight is that the chairman had the intention or desire to help the environment. Both are lacking in Help. Therefore, they judge that the chairman did not help the environment intentionally. Thus, “moral considerations are playing a helpful role in people's underlying competence itself” (Knobe, 2006, p. 226).⁸⁹

2.2 Multiple Core Concepts Explanation

Shaun Nichols and Joseph Ulatowski (2007) think that Knobe's results are best explained by what they call “interpretative diversity.” They used a within-participants design giving participants both Harm and Help. Nichols and Ulatowski replicated Knobe's asymmetry, and they, like Knobe, had a substantial number of participants dissenting from the majority responses. In fact, they found that roughly a third responded 'no' to both Harm and Help, a third responded 'yes' to Harm and Help, and a third responded 'yes' to Harm and 'no' to Help. Participants were also asked to explain their answers. These explanations fell into two categories: either (a) the chairman lacked a desire to bring about the side effect, or (b) the chairman foresaw that the side effect would be brought about. Those who judged the chairman brought about the

⁸⁹ It is not clear that Knobe's account is correct. In similar scenarios, Mark Phelan and Hagop Sarkissian (in press) find that when asked specifically about the badness of the side effect, people do not think the side effect is especially bad, yet they think the chairman brought it about intentionally. Additionally, Steven Sverdlik (2004) finds that if one regretfully brings about a bad side effect, then most people do not think the person does so intentionally. It looks like badness of the side effect does not completely explain the Knobe effect.

side effect intentionally justified their answers by (b). Those who did not think the chairman brought about the side effect intentionally justified their answers with (a). Hence, Nichols and Ulatowski conclude that “considerations of outcome may influence which interpretation the term [intentionally] is given” and those interpretations are reflective of two separate concepts of intentional action—a knowledge based concept and a motive based concept (Nichols & Ulatowski, 2007, p. 360).⁹⁰

Fiery Cushman and Alfred Mele (Mele & Cushman, 2007; Cushman & Mele, in press) also find evidence for multiple folk concepts of intentional action. Using a within participants design, Mele and Cushman gave participants the chairman scenarios. In addition, participants were given scenarios where the person in the scenario (1) has a belief that the side effect will be brought about yet lacks a desire to bring it about, or (2) has a desire to bring about the side effect but lacks a belief that he will bring it about. Because they used additional scenarios featuring (1) and (2), they improve on Nichols and Ulatowski's experiment by filling in the missing conditions. Like Nichols and Ulatowski, they found that there are three patterns of responses—some who answer 'yes' to both Harm and Help, some who answer 'no' to both, and some who answer 'yes' to Harm but 'no' to Help. These results suggest that most people think “an action is intentional if it is performed with desire, given the necessary background conditions (which do not include belief)” (Cushman & Mele, in press). However, people differ on whether belief that a side effect will be brought about is sufficient for a side effect to be brought about intentionally. They conclude that there are at least two concepts of intentional action—one that treats belief as a sufficient condition for acting intentionally (explain answers of 'yes' to Harm and Help) and one that treats desire as a necessary condition for acting intentionally (explains answers of 'no' to Harm and Help).⁹¹ They speculate that there might be a third concept that treats desire as a necessary condition except in morally bad cases, in which case it treats belief as a sufficient condition for acting intentionally (which explains the asymmetric answer).⁹²

90 The asymmetric answer is speculated to be the result of “flexibility” in interpreting 'intentionally'.

91 Cushman and Mele define “belief” and “desire” the following ways: “By “belief” we mean agents’ *justified* beliefs that *they will perform those actions* when they are *very confident* that this is so. By “desire” we mean agents’ desires *to perform those actions* (either as ends or as means to ends)” (in press).

92 Some evidence for the third concept is an order effect they discovered. When Harm is presented after Help, Harm gets a significantly lower intentionality rating than when it comes before. This suggests that people who otherwise would be subject to the Knobe effect are influenced into thinking that belief is not sufficient for performing a bad action intentionally.

The common theme in the constructivists' uses of the Knobe effect is that they use the empirical results to construct (partial) analyses of folk concepts of intentional action. One crucial assumption of these constructivists' accounts is that the responses of the individuals reflect something about the concept of intentional action. If the intuitions generated from the Knobe effect do not tell us anything about folk concepts of intentional action, then it is illegitimate to construct analyses of folk concepts based on those responses. Alternative, critical explanations of the Knobe effect are the focus of the next section.

3. Critical Explanations of the Knobe Effect

3.1 Pragmatic Explanations

Fred Adams and Amie Steadman (2004a, 2004b) think the Knobe effect is best explained by conversational implicature. They argue that in Harm participants want to say that the chairman is blameworthy. The only way they can express this blame is by saying that the chairman brought about the side effect intentionally. If the participants say the chairman did not bring about the bad side effect intentionally, then that would conversationally imply that the chairman is not blameworthy for it. So, participants use “you did that intentionally” language to assign blame to the chairman for bringing about the bad side effect (Adams & Steadman, 2004a, p. 178). However, participants do not want to say the chairman is praiseworthy in Help because he did not care at all about bringing about the good side effect; hence, they judge that the chairman did not help the environment intentionally. Because the results are due to conversational implicature, the judgments made about the chairman cases do not tell us anything about the core folk concept of intentional action (Adams & Steadman, 2004a, p. 178).⁹³

3.2 Biasing Explanations

Nadelhoffer argues the Knobe effect is due to affective biasing. Specifically, Nadelhoffer (2004a) thinks that the perceived blameworthiness of the chairman fuels the asymmetry. Blameworthiness explains the Knobe effect because Harm and Help are not analogous. In Harm and Help, we naturally form negative impressions of the chairman because he does not care about something he should care about. For this reason, people do not want to praise the chairman for bringing about a good side effect, but they do want to blame him for bringing about a bad side effect. Indeed, praise ratings for the chairman intentionally bringing about the good side

⁹³ There are some problems with Adams and Steadman's explanation. First, participants are allowed to express blame, so there is no reason to conversationally imply blame (Malle 2006). Second, in other cases where they presumably want to use the intentionality language to blame, they are not affected by debiasing scenarios where they are shown that a person can be blameworthy and do something unintentionally (Knobe 2003b).

effect are much lower than blame ratings for the chairman bringing about the bad side effect. Because the chairman in Harm is seen as blameworthy and the chairman in Help is seen as blameworthy, the cases are not truly analogous. In order to be truly analogous, the chairman in Help needs to be seen as praiseworthy—something that is not the case in Help. Hence, because in both cases the chairman is seen as blameworthy, participants think that the chairman brings about the harm intentionally but not the help.

This generates the hypothesis that if the cases were truly analogous—if the person in the help condition were perceived as praiseworthy—then people would judge that person brings about the side effect intentionally. Nadelhoffer (2004c) tested this hypothesis. His case describes two friends who are competing against each other in an essay contest. One of the friends, Jason, helps edit his friend's essay, unconcerned that doing so will lower his own chances of winning the contest; and Jason in fact does not win. Most people (55%) judge that Jason decreases his chances intentionally and is praiseworthy for doing so. Therefore, when the cases are analogous, the Knobe effect vanishes (Nadelhoffer, 2004c, p. 210).

Nadelhoffer garners additional evidence that blameworthiness influences intentionality ratings. He uses two cases that have the same evidential features. In one case, a thief attempts to flee from a policeman who jumps onto the thief's car to eventually be shaken off and killed. The thief doesn't care at all about the policeman; he just wants to get away. In the other version, a carjacker jumps onto the hood of the thief's car to be eventually shaken off and killed. The thief does not care at all about the carjacker; he just wants to get away. Thirty-seven percent of participants thought that the thief intentionally killed the policeman, whereas only 10% thought the thief intentionally killed the carjacker. The mean blame rating in the policeman case was 5.11 on a 6 point scale (1 = no blame, 6 = a lot of blame) but only 2.01 for the carjacker (Nadelhoffer, 2006b, p. 209). Therefore, the perception of the target of the side effect influences intentionality ratings even when all other evidential features of the case stay the same. Because all the evidential features are the same, Nadelhoffer thinks these results suggest that when “morally loaded features are built into scenarios, these features often trump or override the standard application of the concept of intentional action—thereby distorting our judgments about intentionality” (2006b, p.213-4).

Bertram Malle and Sarah Nelson (2003) think that negative affect generated in the chairman cases can explain the Knobe effect. They argue that negative affect can bias one's interpretations of the mental states of another person. They studied the judgments of couples who

have fought. Because of the negative affect that is generated in these fights, the parties to the fight think that everything the other person does is intentional—even if it is really not (Malle & Nelson, 2003, p. 575). Likewise, because we think that the chairman has done something wrong, which triggers a negative reaction, our judgments about intentionality are biased. Because people see the chairman in a negative light, they are more likely to think that he brings about the harmful side effect intentionally.

3.3 The Attention Explanation

Malle (2006) argues that the way in which the scenarios are presented focuses the participants' attention to the evaluative components of the scenarios. For example, the chairman does not care at all about the bad side effect. The participants may think that they are supposed to “do” something with this evaluative material, especially because the side effect is so extreme. Because they think they are supposed to do something with this evaluative material, they use it to make “non-technical” intentionality judgments. That is, because they are forced to make intentionality judgments about the chairman, they use this information to determine that the chairman brought about the side effect intentionally. However, in other contexts where they do not think they are supposed to use this material, the participants would be more inclined to use their core concept and issue a judgment that the chairman did not bring about the bad side effect intentionally.⁹⁴

The common thread of the critical explanations of the Knobe effect is that they all entail that the Knobe effect does not tell us anything about folk concepts of intentional action. That is, the Knobe effect is explainable without reference to the core concept of intentional action. If the results of the Knobe effect do not inform us of folk concepts of intentional action, then these results do not tell us in what ways our philosophical analyses about intentional action are supposed to be constrained.

4. Preliminary Glance at Folk Intuitions about Intentional Action

There are a couple constructivist explanations of the Knobe effect suggesting that the Knobe effect may reflect one or more core concepts of intentional action (Cushman & Mele, in press; Knobe, 2006; Nichols & Ulatowski, 2007). However, the critical explanations suggest folk intuitions are too unreliable to be used in philosophical analyses. First, there are explanations of the intuitions that do not reference the folk concept of intentional action. Second,

⁹⁴ Malle does not provide data for this interpretation. However, given the other data on blaming, it makes for a parsimonious explanation.

we have reviewed several examples where non-evidential features of the cases seem to change people's intentionality judgments. For example, Cushman and Mele (in press) report that there are order effects present on people's intentionality judgments. The order of the scenarios is not a property of the side effect described in the scenarios. Because the order of the scenarios is not a property of the side effect, it must be that people are sensitive to extraneous factors. If the order of presentation is an extraneous factor that influences intuitions, then we have reason to think the chairman cases do not indicate the contours of folk concepts of intentional action. If that is so, then we should be apprehensive that the subjects' responses about side effects are indicative of their concept of intentional action. If the critics are right, then there are two problems for the constructivist: (1) folk intuitions may be too unreliable to be used as evidence or adequacy constraints theories and analyses, and (2) philosophical intuitions may be equally subject to extraneous features of cases, making them unreliable as well. To assume that philosophical intuitions are reliable is simply to beg the question against those who think that the unreliability of folk intuitions may generalize to philosophical intuitions about intentional actions.

These data raise several interesting questions. Are the critics right that intuitions do not tell us anything about folk concepts of intentional action? Are they right that we should be wary that philosophical intuitions may be just as unreliable as folk intuitions? Are there stable intuitions about intentional actions? How can we decide if the critics or the constructivists are right about intentional action? Answering these questions is the task of the next two sections.

5. More Data on Folk Intuitions about Intentional Actions

Given that there is evidence that there are different folk concepts of intentional action (cf., Cushman & Mele, forthcoming; Nichols & Ulatowski, 2007), it is plausible that individual differences can help explain, in part, these differing results. After all, there is a great deal of evidence that individual differences can account for differences in reasoning processes (Hyde, 1990). If we can identify groups of people who answer as they do, then we can begin to understand the judgment processes of these groups. We can then start to answer the reliability question—perhaps some groups of people's intuitions are reliable and others aren't. In order to do just that, Edward Cokely and I ran a series of experiments attempting to identify groups of people who (1) account for the order effect, (2) account for the asymmetry, and (3) can help explain the multiple concepts view.

5.1 Order Effects and Sex Differences

One problematic result for the constructivist involves an order effect in intentional action

ascriptions. Given that Cushman and Mele use the order effect to tease out a possible third concept of intentional action, it is probable that individual differences account for the order effect. In an attempt to understand the order effect, I ran a study with Edward Cokely (Feltz & Cokely, 2007). The first step was to replicate the Knobe effect and the order effect found in Cushman and Mele (in press). Participants were volunteers from introductory philosophy classes at Florida State University. There were two groups of participants. One group ($N=35$) received the scenarios in the Help-Harm order, and the other group ($N=33$) received the scenarios in the Harm-Help order.⁹⁵ The participants were instructed to indicate their opinion about the following two statements corresponding to the relevant scenario: (a) The chairman intentionally harmed the environment; (b) The chairman intentionally helped the environment. The participants were given a 7 point scale, with 'Agree' = 7, 'No Opinion' = 4, and 'Disagree' = 1.

The mean responses to the statements are illustrated in Table 11:

Table 11: Responses to Harm and Help by Order

| | The chairman intentionally helped the environment. | The chairman intentionally harmed the environment. |
|-------------|--|--|
| Help First | 1.89 | 3.8 |
| Help Second | 2.67 | 5.0 |

A one-factor repeated measures ANOVA revealed a large and reliable Knobe effect.⁹⁶ However, the order of the scenarios also affected the responses. Specifically, planned analysis revealed a significant order effect with the responses to Harm,⁹⁷ that did not affect responses to Help.⁹⁸

95 There are 33 participants for the Harm-Help order for two reasons. First, two subjects given this survey version said that their responses should not be used because they either guessed, were too tired, did not read the questions, or randomly selected answers. One subject was excluded because in explaining why he answered the way he did, he directly contradicted his responses. He wrote that the chairman “wanted to intentionally harm the environment” and that he disagreed with the chairman's action and not with the statement provided.

96 $F(1, 66) = 7.06, p < .01, \eta_p^2 = .10$

97 $F(1, 66) = 4.28, p < 0.05$

98 $F(1, 66) = 3.82, p > 0.05$

Hence, our data, like Cushman and Mele's, suggest that those who received Help first are less likely to think the chairman harmed the environment intentionally than those who received Harm first. In fact, those who received Harm first agreed with the statement that the chairman harmed the environment intentionally ($M=5.0$) whereas those who received Help first did not agree with the statement that the chairman harmed the environment intentionally ($M=3.8$).

We wondered if some individual differences would predict who is susceptible to the order effect. A reanalysis of the results of this experiment found that women, and not men, were the ones affected by order. The mean responses are indicated in Tables 12 and 13.

Table 12: Men's Responses to Harm and Help by Order

| | The chairman intentionally helped the environment. | The chairman intentionally harmed the environment. |
|-----------------------|--|--|
| Help Presented First | 1.7 | 4.4 |
| Help Presented Second | 3.8 | 5.0 |

Table 13: Women's Responses to Harm and Help by Order

| | The chairman intentionally helped the environment. | The chairman intentionally harmed the environment. |
|-----------------------|--|--|
| Help Presented First | 2.1 | 3.3 |
| Help Presented Second | 2.6 | 5.1 |

The order effect was only significant for women,⁹⁹ and was non-significant for men.¹⁰⁰ It appears that only women were significantly influenced by the order of presentation.

We conducted a follow-up study to replicate these results. 95 students at Florida State University participated in this experiment for partial course credit. Table 14 represents the

⁹⁹ $F(1, 38) = 5.59, p < 0.05, \eta_p^2 = .12$

¹⁰⁰ $F(1, 26) = 1.71, p > .05, \eta_p^2 = .06$

overall means.

Table 14: Overall Means to Harm and Help by Order

| | The chairman intentionally helped the environment. | The chairman intentionally harmed the environment. |
|-----------------------|--|--|
| Help Presented First | 2.0 | 4.2 |
| Help Presented Second | 2.3 | 5.8 |

A one factor repeated measures ANOVA revealed the expected large asymmetric Knobe effect.¹⁰¹ As well, a moderate sized order effect anomaly was detected indicating significantly higher ratings when Harm preceded Help.¹⁰² Importantly, ANOVA also revealed an order by asymmetry interaction such that when Help came first ($M=2.0$) it differed from Harm ($M=4.2$) less than when Help ($M=2.3$) came after Harm ($M=5.8$).¹⁰³

Again, we found that women and men were influenced differently. Table 15 represents the mean responses of men, and Table 16 represents the mean responses of women.

Table 15: Men's Responses to Harm and Help by Order

| | The chairman intentionally helped the environment. | The chairman intentionally harmed the environment. |
|-----------------------|--|--|
| Help Presented First | 2.1 | 5.3 |
| Help Presented Second | 2.4 | 4.6 |

¹⁰¹ $F(1, 93) = 148.24, p < .01, \eta_p^2 = .61$

¹⁰² $F(1, 93) = 14.36, p < .01, \eta_p^2 = .13$

¹⁰³ $F(1, 93) = 7.71, p < .01, \eta_p^2 = .08$

Table 16: Women's Responses to Harm and Help by Order

| | The chairman intentionally helped the environment. | The chairman intentionally harmed the environment. |
|-----------------------|--|--|
| Help Presented First | 2.0 | 3.9 |
| Help Presented Second | 2.3 | 6.1 |

The order effect was significant for women,¹⁰⁴ and it was non-significant for men.¹⁰⁵ Thus, it appears that only women were significantly affected by the order of presentation.

5.2 The Knobe Effect and Extraversion

Given that women and men are affected differently by the order of presentation, and that there is evidence that there are gender related differences among men and women (Hyde, 1990), it is plausible that other individual differences may partially explain the Knobe effect. To explore this possibility, we gave participants a battery of individual difference measures in the replication study. We conducted a stepwise linear regression with the Knobe asymmetry as the dependent variable and the Big Five Personality inventory, brief self-control¹⁰⁶, OSPAN¹⁰⁷, SAT¹⁰⁸, and sex¹⁰⁹ as independent variables. Analysis revealed a significant effect of only one variable, extraversion.¹¹⁰ Next, a hierarchical regression was constructed with the same dependent variables and with order (to control for the known order effect), followed by extraversion as an independent variable. The full model was a significant predictor of the Knobe effect.¹¹¹

However, controlling for the order-effect, extraversion continued to predict unique variance.¹¹²

104 $F(1, 70) = 13.86, p < .001, \eta^2 = .17$

105 $F(1, 19) < 1$

106 A three item measure that suggests intuitively wrong answers that takes some self-control to overcome. For example, a bat and a ball costs \$1.10, and the bat costs \$1 more than the ball. How much does the ball cost?

107 A measure of “cognitive span” or how much one can retain in one's working memory.

108 A self-report of participant's SAT scores.

109 Either male or female.

110 $\beta = .29, t = 2.46, p = .02, R^2 = .08$

111 $F(2, 89) = 7.71, p = .001, R^2 = .15$

112 $\beta = .27, t = 2.68, p = .01, R^2_{\text{change}} = .08$

Given that there is a clear, positive relationship between extraversion and asymmetrical responses, planned analyses next split scores into top and bottom quartiles. ANOVA revealed a clear and reliable extraversion (low, high) by Knobe effect interaction.¹¹³ The Knobe effect was much smaller and qualitatively different for participants who were low in extraversion (Help $M=2.77$, Harm $M=4.37$) as compared to those high in extraversion (Help $M=2.00$, Harm $M=5.79$). Neither sex nor order interacted with the Knobe effect.¹¹⁴ Extraverts were differentially affected by the Knobe manipulation: Individuals who were low in extraversion did not agree that the chairman intentionally harmed the environment, whereas high extraversion scores predicted firm agreement that the environmental harm was intentional—a qualitative difference in these folk judgments.

5.3 The Knobe Effect and Framing Effects

Because the folk are subject to order effects, it seemed likely they would be susceptible to other framing effects. As it turned out, they are. Participants were presented with revised versions of the original Knobe scenarios so the harm or help is described in positive or negative terms. However, the harm and the help that the chairman brings about is *exactly the same*. The two scenarios are:

Negative Chairman: The vice-president of a company went to the chairman of the board and said, “We are thinking of starting one of 2 new programs. Either one will help us increase profits for this year’s balance sheet, but in ten years they will (harm/help) the environment by affecting up to 6,000 acres of a rain forest. If we choose program A, 4,000 acres of the forest will be destroyed with certainty. If we choose program B, there is a 67% chance that 6,000 acres will be destroyed.” The chairman answered, “I don’t care at all about the size of the rain forest. I just want to make as much profit for this year’s balance sheet as I can. Let’s start program A.” They started the new program. Sure enough, in ten years the rain forest was (harmed/helped).

Positive Chairman: The vice-president of a company went to the chairman of the board and said, “We are thinking of starting one of 2 new programs. Either one will help us increase profits for this year’s balance sheet, but in ten years they will (harm/help) the environment by affecting up to 6,000 acres of a rain forest. If we choose program A, 2,000 acres will be saved with certainty. If we choose program B, there is a 33% chance

113 $F(1, 36) = 7.65, p < .01, \eta_p^2 = .18$

114 ($F < 1$)

that 6,000 acres will be saved.” The chairman answered, “I don’t care at all about the size of the rain forest. I just want to make as much profit for this year’s balance sheet as I can. Let’s start program A.” They started the new program. Sure enough, in ten years the rain forest was (harmed/helped).

Participants were volunteers from lower level philosophy classes. Participants were asked to what extent they agreed with one of the following relevant statements: the chairman intentionally (harmed/helped) the environment (1=strongly agree, 4=neutral, 7=strongly disagree). The only difference between the two cases is the valence of the side effect and whether or not the harm or help is described positively or negatively.¹¹⁵

Participants were 127 undergraduates at Florida State University taking introductory-level philosophy courses. Participants received two scenarios. They were given either the Help version framed positively and negatively or the Harm version framed positively and negatively. Because there were no significant within-participant effects, we only analyzed the first response participants gave. Again, there was an overall Knobe effect where those who received the Help version thought that the chairman did not bring about the Help intentionally ($M= 5.8$) whereas those who received the Harm version we neutral about the intentionality of the harm ($M=3.9$).¹¹⁶ However, the frame did influence responses to the Help version. Table 17 displays the mean responses.

Table 17: Responses to Harm and Help by Positive and Negative Frame

| | Intentionally Helped | Intentionally Harmed |
|----------------|----------------------|----------------------|
| Positive Frame | 5.47 (N=29) | 4.14 (N=28) |
| Negative Frame | 6.24 (N=30) | 3.8 (N=30) |

When participants were presented with the negative frame, they strongly disagreed that the

¹¹⁵ One might worry that in the Positive chairman, help condition we assume that subjects infer that 4,000 acres will be destroyed in order to make it analogous to the Negative chairman, help condition. Participants may not make this inference.

¹¹⁶ $t(1, 115) = 6.306, p < 0.01$. Of note, the frame did reduce the severity of the Knobe effect asymmetry.

chairman brought about the help intentionally (M= 6.24). When participants were presented with the positive frame, they moderately disagreed that the help was brought about intentionally (M= 5.47), a significant difference.¹¹⁷ Hence, the way Help was presented influenced judgments. When the help scenario was framed positively, the chairman was more likely to be judged to bring about this side effect intentionally.

6. The Relevance of Experimental Results for Traditional Action Theory

So far, we have found philosophers using data about folk intuitions about intentional action to argue for one folk concept or multiple folk concepts. We have also found evidence that the folk are susceptible to a range of framing effects. These effects suggest that folk intentional action intuitions are not reliable. If an analysis of intentional action that is not constrained by the folk concept is nothing but a philosophical fiction, then the empirical results have some bearing on traditional action theory. But what conclusions can we draw, and how important are they? In what follows, I argue that the critics are correct that we should be cautious when using both folk and philosophical intuitions. However, once the necessary precautions are taken, we can draw some conclusions about the reliability of intuitions.

6.1 Critical Implications

First, philosophers cannot assume that their view fits with the folk view of intentional action. As Nadelhoffer (2006a) argues, the results of the Knobe effect are contrary to what would be predicted by some proponents of the SV. Given the experimental evidence, it appears that the SV does not comport with the folk view of intentional action.¹¹⁸ In Harm, most people think the chairman intentionally brings about the side effect even though he does not intend to do so—a result contrary to the SV. To the extent that we should accept the SV because it comports with the everyday concept of intentional action, it is unmotivated. These philosophers are simply wrong about what the folk intuitions are.

Given the results discussed in this chapter, there may be a more general criticism of the reliability of intuitions about intentional action. We have discovered what looks to be irrational answers to questions about intentional action. On some views of intentional action (e.g., the Simple View), the Knobe effect itself may be considered an instance of irrationality because the relevant *mental* states of the chairman are the same in Harm and Help. The mental items that

¹¹⁷ $t(1, 57) = 2.260, p < 0.05$.

¹¹⁸ I will argue that there is no “the” folk concept or view of intentional action. However, even if there was just one single concept of intentional action, the majority of folk do not respond in ways predicted by a proponent of the SV.

constitute the chairman's decision to implement the plan are the same—he only wants to make money, and he doesn't care at all about the side effect. Considerations of the side effect do not even partly constitute part of the reason that leads him to decide as he does. If the mental states are the same and mental states are what determine whether an event is intentional, then it is irrational to think that the chairman intentionally brought about the harm but did not intentionally bring about the help. Of course, this does not mean that these features do not play a role in people's intuitions about intentional actions. It simply means that, on some views, it is irrational to make intentionality judgments based on extra-mental or irrelevant considerations. If the folk are influenced by extra-mental or irrelevant factors, then we should be cautious interpreting these results as indicating the folk concept of intentional action.

There is additional evidence that the folk are influenced by non-evidential, extra-mental factors. First, there is evidence for an order effect in the chairman cases. Participants are more likely to say that the harm is brought about intentionally when it is presented first than when it is presented after Help. The order of presentation is not an evidential feature of the cases, so it seems irrational to change answers in response to the order of presentation. That is, the order of presentation should not have a “slot” in a concept of intentional action because the order of presentation is a non-evidential feature of the cases. If the order does have an effect, then the reliability of people's intentional action intuitions is questionable.

Second, there is a framing effect for responses to Help. While this effect does not qualitatively affect responses, it does significantly alter the responses. Specifically, when participants are presented with Help framed positively, they disagree that it was done intentionally less than when the case is framed negatively. This framing effect provides additional evidence that people's intentional action intuitions are influenced by non-evidential features of the cases. The exact same side effect is brought about—namely, 2000 trees exist that otherwise might not have existed. Simply framing that side effect differently should not change the intentionality judgments of that side effect. Therefore, if participants are subject to framing effects, we have even less reason to think that their intuitions reliably tell us anything about the concept of intentional action.

There are two upshots of these seemingly irrational intuitions. First, to test proposed analyses of intentional actions we appeal to cases. The worry is that philosophers might also be influenced by non-evidential features. Often our intuitions are biased in ways of which we are not consciously aware. As I argue in Chapter 1, we cannot assume from the outset that

philosophical intuitions are immune from these types of biases. Second, it is unclear our philosophical analyses are properly constrained if we aren't sure what intuitions indicate our concept of intentional action. Again, if the arguments of Chapter 1 are correct, we cannot simply assume that philosophers know the appropriate ways analyses of intentional actions are to be constrained.

To summarize, there are three major critical implications of these results. The first is that philosophers cannot assume, a priori, that what they take to be the intuitions of the folk are indeed the intuitions most folk have. This is so because it appears that some philosophers who have thought that their view is in line with folk intuitions have been wrong. Second, folk intuitions appear to be too unreliable to be evidence for analyses of intentional action. We have seen that the folk are sensitive to a number of non-evidential features of cases that are seemingly irrelevant to the concept of intentional action. If people are influenced by these non-evidential features, then we have good reason to suspect that their intuitions are too unreliable to tell us anything important about their underlying concepts. Third, philosophers who make explicit appeal to their own intuitions may be similarly influenced, possibly unwittingly, by non-evidential features of cases. At a minimum, it would be question begging for philosophers to assume otherwise given the impressive list of biases to which people are subject. If these philosophers are to claim that their intuitions are not unstable, then they must provide some empirical evidence to that effect.

6.2 Hopes for the Constructivist

There is a strong case to be made that intuitions about intentional actions are not reliable. The impressive list of biases and framing effects indicate that people are not sensitive to only the evidential features of the cases. These effects suggest that folk intuitions, and by proxy, philosophical intuitions may not be reliable enough to be used as evidence for philosophical analyses of intentional action. Is there hope for the constructivist to answer the problems presented by the critics? I believe that there is.

There are two general ways for the constructivist to respond to the data suggesting that our intuitions about intentional actions are not reliable. First, they could attempt to explain and provide additional data that the seeming instability of intuitions is in fact a reflection of different sub-groups of people responding in different ways. That is, they may attempt to show that there are groups who express stable intuitions, and those stable intuitions are reflective of different concepts. If we find that there are different groups who express different intuitions, then we have

good reason to think that while the critics are right and there is no one stable concept of intentional action (because there simply is no single concept), we can offer different analyses of different folk concepts. Our philosophical analyses would then be constrained by those concepts.

Second, there could be really just one concept of intentional action, and the apparent instability of intuitions is due to the fact that different people interpret scenarios differently. That is, some people may interpret the scenarios as asking about their non-technical notions of intentionality. If people interpret scenarios differently, then there will be apparent instability in their intuitions because different scenarios will be interpreted differently by different people. As previously discussed, there is a body of literature that suggests some groups of people code and retrieve data differently from some other groups. If this is true, then it is unclear that either (1) the responses to the scenarios indicate something about their concept of intentional action, or (2) that their intuitions are unstable. It is possible that if people were to interpret the scenarios the same, they might have the same intuitions about the cases.

The results reported in this chapter support the view that there are different groups of people who express stable intuitions. If these results are right, then we can identify the groups of people who give the pattern of responses that they do. We can identify women as those responsible for one seeming instability of intuitions—the order effect. Because we know there are gender based differences in reasoning, concepts, and memory (Hyde, 1990), we have reason to think that women simply have a different concept of intentional action than men. Following Cushman and Mele (in press), we can infer that women normally think that foresight isn't sufficient for acting intentionally except for morally bad actions.¹¹⁹ Of course, this leaves open exactly who has the other two concepts of intentional action, but it does identify one group and offers us a way to explain what was a seeming instability in intuitions.¹²⁰

Additionally, the data indicate that people who report themselves to be extraverted are responsible for the asymmetric response indicative of the Knobe effect. Participants who scored low on extraversion tend to respond the same way to Harm and Help. We also know that extraverts have different judgment processes in other domains. For example, extraversion is

¹¹⁹ Cushman and Mele (in press) do not discuss gender differences.

¹²⁰ It is not clear why women are subject to an order effect about intentional action intuitions and men are subject to an order effect for epistemological intuitions. One possibility is that women are more socially sensitive than men. Women may be more likely to think that chairmen have special obligations because of their position of power whereas men do not. That could help explain why women, and not men, are subject to order effects on their intentional action intuitions.

negatively correlated with general intelligence (Chamorro-Premuzic, Furnham, & Ackerman, 2006), extraverts tend to code information differently (Rusting, 1999; Szalma & Hancock, 2005), and they make different predictions about the future (Zelenski & Larsen, 2002). Therefore, because extraverts are responsible in large part for the Knobe effect, we have good reason to think that extraverts have a different concept of intentional action from non-extraverts.¹²¹

However, these results suggest an alternative hypothesis. People may simply interpret the scenarios differently. Because we have found that women are largely responsible for the order effect, and extraverts are largely responsible for the Knobe effect, it could be that these groups of people simply interpret the scenarios differently. Because they interpret the scenarios differently, they give a non-technical judgment about the cases. For example, there is good evidence that women are particularly sensitive to harms to the environment (Wilson et al., 1996). Therefore, women may be more likely to interpret the chairman as a bad man. Seeing the chairman as bad generates negative affect, and this negative affect can lead women to think that the chairman harmed the environment intentionally. Likewise, as noted above, extraverts interpret these scenarios differently from non-extraverts, so it is likely that the valence of the side-effects affect extraverts differently. They may see the harm as one where one in a position of power violates a social convention not to do harm. Because extraverts are more sensitive to the social nature of the harm, they may think that the chairman brought about the harm intentionally. If people only interpret scenarios differently, then there may be just one folk concept and people may apply (or fail to apply) that concept differently depending on their interpretations of the scenarios.

Of course, these hypotheses can neither be confirmed nor disconfirmed with the present data. This calls for more research using more powerful methods. For example, fine grained processing tracing techniques such as talk aloud and protocol analysis may be required to determine exactly how these different groups of people interpret the scenarios (Ericsson & Simon, 1993). This future research would be able to tease out the individual differences among different groups of people and would be able to arbitrate between these two hypotheses.

In any event, we have good reason to think that there is some stability in intentional action intuitions. However, this stability takes a surprising form. Instead of all people expressing the same intuitions about these cases, we find that there are groups definable with stable individual differences that express stable intuitions. Because we know these groups of people

¹²¹This runs somewhat counter to Cushman and Mele's claim that the asymmetric answer can be accounted for by extrapolating what the order effect indicates. Extraversion continued to predict unique variance after the order effect was taken into account. So, perhaps this indicates there is a fourth concept.

have different judgment processes, two possibilities emerge: (a) these groups have different concepts; or (b) these groups have the same concepts yet interpret the scenarios differently. Either of these two possibilities is sufficient to hold the critics at bay. (a) is sufficient because these people are not irrational because they are expressing intuitions in line with their concepts. (b) is sufficient because even though these people interpret the scenarios differently, they have the same underlying core concept.¹²²

7. Concluding Remarks

The results from this chapter suggest three things. First, some care needs to be taken when interpreting folk responses to surveys. There are a variety of ways intuitions can be illegitimately influenced, and those ways need to be taken into account. Second, if the appropriate care is taken, folk intentional action intuitions are reliable enough to be used as evidence for conceptual analyses and theories. Given the prominent role folk intuitions have in theories and analyses of intentional action, it is important to get a clear grasp on these folk intuitions. Otherwise, we won't know what constrains analyses of intentional actions. Finally, we might find that some philosophers' analyses of intentional action might be supported only by some group's intuitions about intentional actions.

¹²² While I do not have any data about individual differences and framing effects, given all the data reported in this dissertation about how various seeming instabilities can be accounted for by groups of people, we have good reason to think that framing effects will be handled in a similar fashion.

CHAPTER 7: THE RELEVANCE OF FOLK INTUITIONS TO PHILOSOPHICAL DEBATES

Nisbett and Wilson (1977) report data that suggest people are often not aware of what influences the choices they make. Not only that, but when people make judgments, they often *make up* justifications for their choices. The critic claims that something similar happens in the case of intuitions. We do not yet know what mechanisms generate intuitions nor are we consciously aware of what influences our intuitions. So we cannot be sure from introspection just how reliable our intuitions are. The critic's challenge is especially severe because, as I have argued, there is no satisfactory a priori defense of intuitions. These a priori attempts to establish intuitions' reliability either (1) beg the question against the critic, (2) are not strong enough to establish a priori that intuitions are reliable without empirical evidence, or (3) do not establish that there are enough philosophically relevant domains where intuitions are reliable. If there are no existent a priori defenses of intuitions in many philosophically relevant domains and if the reliability of these intuitions is to be established, it must be done empirically.

Given that there is no adequate a priori defense of the reliability of intuitions in philosophy, the reliability of intuitions is in part an empirical matter. If the reliability of intuitions is in part an empirical matter, then philosophers can have two views about intuitions: (1) intuitions are important to philosophical theories and conceptual analyses or (2) intuitions are not important to philosophical theories. Many philosophers accept (1). Those who adopt (1) use folk intuitions in at least one of two ways: (a) folk intuitions are used as evidence for philosophical theory and analyses or (b) folk intuitions are used as adequacy constraints on theories and conceptual analyses. Because folk intuitions are used in these ways, these theories are in part dependent on what the folk think about philosophically important questions. Hence, some philosophers who think that (a) or (b) is important give folk intuitions a critical role in philosophical theories and analyses.

When the critic looks at the empirical data, she is not encouraged that they establish the reliability of intuitions. In fact, the data further call into question intuitions' reliability. These data are suggestive of three things that further call into question the reliability of intuitions. First, as we have seen, philosophers from a wide variety of disciplines make reference to folk intuitions. Sometimes, these philosophers are wrong about what the folk intuitions are, but they are wrong

in a unique way. Not only do they misdescribe folk intuitions, philosophers also sometimes claim that they know what features of cases influence folk intuitions. But we have evidence that philosophers are sometimes wrong on both accounts.

Second, if philosophers are poor at understanding what the folk intuitions are and what influences those intuitions, then we have some reason to worry about the reliability of philosophical intuitions as well. As the data of Nisbett and Wilson suggest, we are often not very good at consciously determining what influences our choices and we make up justifications for those choices. Likewise, we do not have conscious access to what generates our intuitions. Because we are not aware of what produces our intuitions, we should be cautious about using intuitions as evidence. It could be that our intuitions, like our choices, are entities that we attempt to justify after we have them and not entities that point out, in any significant sense, the truth about our concepts or how good our theories are.

Third, the problem is compounded by the fact that intuitions seem to be unreliable in a variety of contexts. They are subject to order and framing effects. These effects indicate that intuitions are influenced by extraneous factors. If intuitions are influenced by these extraneous factors, then we have good reason to think that they are not reliable enough for us to use them in philosophical theories or analyses of philosophical concepts.

As the data I have reported suggest, these three worries cut a wide swath in philosophy. We have found seeming instability of intuitions in ethics, epistemology, free will, and action theory. This is a substantial amount of instability in four key areas of contemporary philosophy. Given such instability, the critic might think it justified to argue that intuitions are generally unreliable. After all, there is no satisfactory a priori defense of the reliability of philosophical intuitions, and a substantial amount of evidence points to intuitions' unreliability.

The guiding question of this dissertation is: Why should philosophers care about folk intuitions? According to the critic, philosophers should care about folk intuitions because they indicate several worrisome features. First, some philosophers use folk intuitions to support one theory or analysis of a concept. To the extent that one's theory fits with commonsense intuitions, those theories and conceptual analyses are correct. But the critic is quick to point out that these philosophers need to be *more* concerned with folk intuitions. They have to provide the requisite empirical evidence to use folk intuitions as evidence. The critic notes that in the four domains surveyed in this dissertation, philosophers are sometimes very bad at describing what the folk intuitions are. This indicates that *all* philosophers who take folk intuitions seriously must provide

evidence that what they take to be the intuitions the folk have are the intuitions the folk have.¹²³

However, some philosophers may take the view towards intuitions indicated in (2) and reject the idea that intuitions have any role to play in philosophy. If folk intuitions are unreliable, it would be unwise to use intuitions for evidence or constraints on philosophical theories and conceptual analyses. If folk intuitions are susceptible to biases and errors, then it is not clear for what they provide evidence. Therefore, we shouldn't use folk intuitions to constrain or provide evidence for philosophical theories and analyses. However, taking the view indicated in (2) has some costs of its own. First, there is the risk that theories and conceptual analyses that are not at least in part grounded in everyday notions run the risk of being analyses of nothing more than “philosophical fictions” (Mele, 2001, p. 27). That is, these theories and analyses that do not use everyday intuitions for support or adequacy constraints may not apply to any folk concepts.

Second, the critic argues, these philosophers are not totally immune from empirical data. If philosophers use intuitions to support their theories, even if they are more refined philosophical intuitions, it could be that philosophical intuitions are not reliable. That is, given the evidence of Nisbett and Wilson and the impressive number of biases that folk intuitions are prone to, we should also be apprehensive that philosophical intuitions are any more reliable. After all, philosophers do not have introspective access to what generates or influences their intuitions just as most people do not have introspective access to what influences their intuitions. Given that philosophers don't have introspective access to what generates or influences their intuitions and folk intuitions have been shown to be unreliable, we have reason to think that philosophical intuitions could also be unreliable. At a minimum, philosophers who rely on intuitions should provide some empirical evidence that they are not so influenced.

Constructivists give a slightly different answer to the guiding question of this dissertation, but they agree with the critics in spirit. Constructivists agree with the critics that if philosophers are going to get any theoretical mileage out of folk intuitions, then they must provide some evidence that the intuitions they attribute to the folk are indeed the intuitions that the folk have. Without those data, and with the data indicating that philosophers are often wrong about what intuitions the folk have, constructivists think that gathering data about folk intuitions is critical for their incorporation into philosophical theory and conceptual analyses.

However, the constructivist finds hope that there is enough stability of intuitions to be

¹²³ Of course, there may be some philosophers who are particularly sensitive and can accurately describe folk intuitions. If there are such philosophers, an exception can be made for them. But, if there are any, they are small in number and I have not met any.

used in philosophical theories and conceptual analyses. In each of the philosophical domains discussed in this dissertation, there is apparent evidence of instability in folk intuitions. However, in each of these domains I have argued that there is an underlying stability in folk intuitions. This stability expresses itself in individual groups of people having different, stable sets of intuitions.

This group level stability is sufficient to warrant using intuitions for philosophical theory and conceptual analyses. Philosophers who accept (1) can take solace in the fact that there is some stability in intuitions. But it is important to distinguish two theses that could make up part of (1):

The Unity of Folk Intuitions (UFI): in order for folk intuitions to be evidence for a theory or analysis (i.e., they are philosophically important), then (i) all (or most) folk intuitions must (ii) reliably indicate something.

The Fragmentation of Folk Intuitions (FFI): in order for folk intuitions to be evidence for a theory or analysis (i.e., they are philosophically important), then (i) some folk intuitions must (ii) reliably indicate something.

The evidence suggests that UFI is false. First, the evidence of framing effects and biases indicates that folk intuitions taken as a whole, then they are responsive to extraneous factors. Intuitions that are responsive to extraneous features are not reliable, making (i) of UFI false. Therefore, UFI cannot be the correct thesis. However, the evidence suggests that FFI is true. There are groups of people who appear to have stable intuitions about philosophically important cases across the four philosophically relevant domains where intuitions have been called into question. Some intuitions are not subject to order effects or other biases. Because some people's intuitions are not subject to order effects and biases, parts (i) and (ii) are satisfied. Therefore, FFI appears to be true.

Because FFI appears to be true, the constructivist argues that intuitions are relevant and can be used provide evidence and constrain philosophical theories and analyses. The troubling result the critic calls upon to undermine the use of folk intuitions in philosophy is folk-wide instability of intuitions. But because there is evidence that there is sub-group stability in the sense that these groups are not influenced by extraneous factors, we do not have reason to doubt the reliability of those intuitions. So, we have the requisite empirical evidence to support the constructivist's claim that folk intuitions can be used as evidence and can constrain philosophical theories and analyses.

Of course, there may be problems with FFI. One obvious question to ask is “How

fragmented can folk intuitions be and still offer evidence or constrain philosophical theories and analyses?" Can a group consist of a single person S whose intuitions reliably indicate something? Given that S's intuition reliably indicates something, FFI is satisfied. S's intuition can be used as evidence or to constrain philosophical theories and analyses. But, S may have very strange intuitions. FFI would appear to license any philosophical theory using anyone's intuitions just as long as those intuitions reliably indicate something. That might have the strange result of a multitude of acceptable theories and analyses that seem wrong to the majority of people and lead to very unpleasant results (e.g., the psychopath may have the intuition that it is permissible to kill people at will).

On one level the objection is correct. If FFI is true, then any intuitions can be used as evidence just in case they reliably indicate something. But that shouldn't strike one as odd. After all, there are many examples of idiosyncratic intuitions that philosophers have that are meant to provide evidence for philosophical theories or analyses. For example, Kirk Ludwig thinks that it is possible to intend to do something that one believes is impossible (1992). Very few (if any) people share Ludwig's intuitions about the cases he presents.

Even if the objection is correct insofar as it goes, the seeming force of the objection can be deflated. It is an open empirical question whether folk intuitions are actually so fragmented. The objection trades on the *possibility* of intuitions being objectionably fragmented. But in order for the objection to have any real bite, it must be shown that intuitions are actually fragmented in an objectionable way. While there is not at present enough evidence to settle this issue, the evidence is suggestive of wide agreement in groups of people. If there is wide agreement of intuitions, then FFI may not actually lead to the objectionable result.

Even if FFI would lead to the objectionable result that folk intuitions provide evidence for things that seem to be obviously false, it is an error to think that *just describing* folk intuitions is going to *settle* philosophical issues. The constructivist is open to the possibility that some folk intuitions are wrong. She is also open to the possibility that analyses and theories that use folk intuitions as evidence are false. However, showing that these theories are false requires additional argument. So, while folk intuitions provide evidence for theories, that does not entail that the theories that are based on that evidence are necessarily true. The type of evidence intuitions provide is defeasible, but evidence nonetheless.

In sum, the evidence I have presented provides a strong presumption in favor of FFI and the constructivist's view of folk intuitions. Of course these are first steps to understanding a number

of issues involving folk intuitions. We should also want to know what processes generate folk intuitions, under what conditions folk intuitions are likely to err, and in what domains we get fragmentation of intuitions. But, when the proper care is taken, intuitions are found to be reliable enough to provide evidence for and constrain philosophical theories and analyses. Therefore, folk intuitions are relevant to philosophical debates.

APPENDIX A

IRB APPROVAL AND SAMPLE VERBAL CONSENT FORM

Date: 1/26/2007

To: Adam Feltz

Address: 1500

Dept.: PHILOSOPHY DEPARTMENT

From: Thomas L. Jacobson, Chair

Re: Use of Human Subjects in Research
Folk Intuitions, Individual Differences, and Philosophical Questions

The forms that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be Expedited per 45 CFR § 46.110(7) and has been approved by an expedited review process.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.

If you submitted a proposed consent form with your application, the approved stamped consent form is attached to this approval notice. Only the stamped version of the consent form may be used in recruiting research subjects.

If the project has not been completed by 1/25/2008 you must request a renewal of approval for continuation of the project. As a courtesy, a renewal notice will be sent to you prior to your expiration date; however, it is your responsibility as the Principal Investigator to timely request renewal of your approval from the Committee.

You are advised that any change in protocol for this project must be reviewed and approved by the Committee prior to implementation of the proposed change in the protocol. A protocol change/amendment form is required to be submitted for approval by the Committee. In addition, federal regulations require that the Principal Investigator promptly report, in writing any unanticipated problems or adverse events involving risks to research subjects or others.

By copy of this memorandum, the Chair of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Human Research Protection. The

Assurance Number is IRB00000446.

Cc: Alfred Mele, Advisor
HSC No. 2006.219

Verbal Consent Form

I am a graduate student under the supervision of Dr. Alfred Mele in the Department of Philosophy at Florida State University. I am conducting a research study to see what non-professional philosophers think about some philosophically important questions.

I am recruiting subjects to read some scenarios and honestly answer questions about them. This will take approximately 30 minutes.

Your participation in this study is voluntary. If you choose not to participate or to withdraw from the study at any time, there will be no penalty; it will not affect your grade. The results of the research may be published, but your name will not be used. If you choose not to participate, simply turn in a blank survey.

In order to ensure anonymity, please be sure to leave no identifying marks on your survey. This means that you should not write your name, social security number, or anything else that might identify the survey as belonging to you. The confidentiality of any personal information will be protected to the extent allowed by law.

There is minimal risk harm or benefit for you participating in the survey. Participating in the survey carries no more than risk than arises in everyday life. While there is minimal risk of harm or benefit to you, your participation in this survey will benefit the philosophical community by helping shed light on what non-professional philosophers think about philosophically important questions.

If you have any questions concerning this research study, please contact me or Dr. Mele. I can be reached at (850) 644-4127, or email me at adf04@fsu.edu. Dr. Mele can be reached at (850) 644-0217 or by email at almele@mail.fsu.edu. You may also contact the Florida State University Institutional Review Board for more information and your rights as a participant in this survey at their:

(a) Physical Location:

2010 Levy Ave Bldg B Suite 276
Tallahassee, FL 32310

(b) Mailing Address:

Florida State University
Human Subjects Committee
Tallahassee, FL 32306-2742

(c) Email:

jth5898@fsu.edu

APPENDIX B

STANLEY'S SCENARIOS

Low Stakes: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. It is not important that they do so, as they have no impending bills. But as they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Realizing that it isn't very important that their paychecks are deposited right away, Hannah says, 'I know the bank will be open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.'

High Stakes: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. But, as Sarah points out, banks do change their hours. Hannah says, 'I guess you're right. I don't know that the bank will be open tomorrow.'

Ignorant High Stakes: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. But neither Hannah nor Sarah is aware of the impending bill, nor of the paucity of available funds. Looking at the lines, Hannah says to Sarah, 'I know the bank will be open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.'

Low Attributer-High Subject Stakes: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Two weeks earlier, on a Saturday, Hannah went to the bank, where Jill saw her. Sarah points out to Hannah that banks do change their hours. Hannah utters, 'That's a good point. I guess I don't really know that the bank will be open on Saturday.' Coincidentally, Jill is thinking of going to the bank on Saturday, just for fun, to see if she meets Hannah there. Nothing is at stake for Jill, and she knows nothing of Hannah's situation. Wondering whether Hannah will be there, Jill utters to a friend, 'Well, Hannah was at the bank

two weeks ago on a Saturday. So she knows the bank will be open on Saturday'

High Attributer-Low Subject Stakes: Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday Hannah calls up Bill on her cell phone, and asks Bill whether the bank will be open on Saturday. Bill replies by telling Hannah, 'Well, I was there two weeks ago on a Saturday, and it was open.' After reporting the discussion to Sarah, Hannah concludes that, since banks do occasionally change their hours, 'Bill doesn't really know that the bank will be open on Saturday'. (Stanley 2005, 3-5)

APPENDIX C

EXPERIMENT WITH STANELY'S SCENARIOS

LS

Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. It is not important that they do so, as they have no impending bills. But as they drive past the bank, they notice that the lines inside are very long, as they often are on Friday afternoons. Realizing that it isn't very important that their paychecks are deposited right away, Hannah says, 'I know the bank will be open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.'

Assume that the bank really will be open tomorrow. Please indicate how strongly you agree with the following statement:

When Hannah says, "I know the bank will be open tomorrow," what she says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

HS

Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. But, as Sarah points out, banks do change their hours. Hannah says, 'I guess you're right. I don't know that the bank will be open tomorrow.'

Assume that the bank really will be open tomorrow. Please indicate how strongly you agree with the following statement:

When Hannah says, "I don't know that the bank will be open on tomorrow," what she says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

IHS

Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. But neither Hannah nor Sarah is aware of the impending bill, nor of the paucity of available funds. Looking at the lines, Hannah says to Sarah, 'I know the bank will be open tomorrow, since I was

there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.'

Assume that the bank really will be open tomorrow. Please indicate how strongly you agree with the following statement:

When Hannah says, "I know the bank will be open tomorrow," what she says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

LAHSS

Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Two weeks earlier, on a Saturday, Hannah went to the bank, where Jill saw her. Sarah points out to Hannah that banks do change their hours. Hannah utters, 'That's a good point. I guess I don't really know that the bank will be open on Saturday.' Coincidentally, Jill is thinking of going to the bank on Saturday, just for fun, to see if she meets Hannah there. Nothing is at stake for Jill, and she knows nothing of Hannah's situation. Wondering whether Hannah will be there, Jill utters to a friend, 'Well, Hannah was at the bank two weeks ago on a Saturday. So she knows the bank will be open on Saturday'

Assume that the bank really will be open tomorrow. Please indicate how strongly you agree with the following statement:

When Jill says, "she knows the bank will be open on Saturday," what she says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

APPENDIX D
MINIMAL AND ATTRIBUTER CASES

LS/NA

Bill, Jim, and Sarah are hiking and they come to a ravine. There is a bridge five feet over the ravine. Bill sees Sarah and Jim cross the bridge, and Bill says to Jim, "I know that the bridge is stable enough to hold my weight."

Assume that the bridge is stable enough hold Bill's weight. Please indicate how strongly you agree with the following statement:

When Bill says, "I know that the bridge is stable enough to hold my weight," what he says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

HS

Bill, Jim, and Sarah are hiking and they come to a ravine. There is a bridge one hundred feet over the ravine. Bill sees Sarah and Jim cross the bridge, and Bill says to Jim, "I know that the bridge is stable enough to hold my weight."

Assume that the bridge is stable enough hold Bill's weight. Please indicate how strongly you agree with the following statement:

When Bill says, "I know that the bridge is stable enough to hold my weight," what he says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

A

Bill, Jim, and Sarah are hiking and they come to a ravine. There is a bridge five feet over the ravine. Bill sees Jim and Sarah cross the bridge, and Jim says to Sarah, "Bill knows that the bridge is stable enough to hold his weight."

Assume that the bridge is stable enough hold Bill's weight. Please indicate how strongly you agree with the following statement:

When Jim says, "Bill knows that the bridge is stable enough to hold his weight," what he says is true.

APPENDIX E
SIMPLIFIED CASES

HS

Hannah and her sister Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, it is very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. Hannah says to Sarah, 'I know that the bank will be open tomorrow.'

Assume that the bank really will be open tomorrow. Please indicate how strongly you agree with the following statement:

When Hannah says, "I know that the bank will be open on tomorrow," what she says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

LS

Hannah and her sister Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they do not have an impending bill coming due, it is not very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. Hannah says to Sarah, 'I know that the bank will be open tomorrow.'

Assume that the bank really will be open tomorrow. Please indicate how strongly you agree with the following statement:

When Hannah says, "I know that the bank will be open tomorrow," what she says is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

APPENDIX F
BRIDGE CASES

LS

John is driving a truck along a dirt road in a caravan of trucks. He comes across what looks like a rickety wooden bridge over a three foot ditch. He radios ahead to find out whether other trucks have made it safely over. He is told that all 15 trucks in the caravan made it over without a problem. John reasons that if they made it over, he will make it over as well. So, he thinks to himself, "I know that my truck will make it across the bridge."

Assume that the bridge is safe enough for him to cross. Please indicate how much you agree with the following statement:

When John thinks to himself, "I know that my truck will make it across the bridge," what he thinks is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

HS

John is driving a truck along a dirt road in a caravan of trucks. He comes across what looks like a rickety wooden bridge over a yawning thousand foot drop. He radios ahead to find out whether other trucks have made it safely over. He is told that all 15 trucks in the caravan made it over without a problem. John reasons that if they made it over, he will make it over as well. So, he thinks to himself, "I know that my truck will make it across the bridge."

Assume that the bridge is safe enough for him to cross. Please indicate how much you agree with the following statement:

When John thinks to himself, "I know that my truck will make it across the bridge," what he thinks is true.

| | | | | | | |
|-------------------|---|---|---------|---|---|----------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Agree | | | Neutral | | | Strongly Disagree |

APPENDIX G

SAW'S SCENARIOS

We are investigating what different people's opinions are about knowledge. In each question, please indicate to what extent you agree or disagree with that statement.

1. Suzy looks out the window of her car and sees a barn near the road, and so she comes to believe that there's a barn near the road. However, Suzy doesn't realize that the countryside she is driving through is currently being used as the set of a film, and that the set designers have constructed many fake barn facades in this area that look as though they are real barns. In fact, Suzy is looking at the only real barn in the area.

Please indicate to what extent you agree or disagree with the following claim: "Suzy knows there is a barn near the road."

Strongly agree Agree Neutral Disagree Strongly disagree

2. Dave likes to play a game with flipping a coin. He sometimes gets a "special feeling" that the next flip will come out heads. When he gets this "special feeling", he is right about half the time, and wrong about half the time. Just before the next flip, Dave gets that "special feeling", and the feeling leads him to believe that the coin will land heads. He flips the coin, and it does land heads.

Please indicate to what extent you agree or disagree with the following claim: "Dave knew that the coin was going to land heads."

Strongly agree Agree Neutral Disagree Strongly disagree

3. One day Charles was knocked out by a falling rock; as a result his brain was "rewired" so that he is always right whenever he estimates the temperature where he is. Charles is unaware that his brain has been altered in this way. A few weeks later, this brain rewiring leads him to believe that it is 71 degrees in his room. Apart from his estimation, he has no other reasons to think that it is 71 degrees. In fact, it is 71 degrees.

Please indicate to what extent you agree or disagree with the following claim: "Charles knows that it is 71 degrees in his room."

Strongly agree Agree Neutral Disagree Strongly disagree

4. Karen is a distinguished professor of chemistry. This morning, she read an article in a leading scientific journal that mixing two common floor disinfectants, Cleano Plus and Washaway, will create a poisonous gas that is deadly to humans. In fact, the article is correct: mixing the two products does create a poisonous gas. At noon, Karen sees a janitor mixing Cleano Plus and Washaway and yells to him, "Get away! Mixing those two products creates a poisonous gas!"

Please indicate to what extent you agree or disagree with the following claim: “Karen knows that mixing these two products creates a poisonous gas.”

___ Strongly agree ___ Agree ___ Neutral ___ Disagree ___ Strongly disagree

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half the lake? _____ days

SWB

Using the 1 - 7 scale below, indicate your agreement with each item

by placing the appropriate number on the line preceding that item. Please be open and honest in your responding.

7 - Strongly agree

6 - Agree

5 - Slightly agree

4 - Neither agree nor disagree

3 - Slightly disagree

2 - Disagree

1 - Strongly disagree

_____ In most ways my life is close to my ideal.

_____ The conditions of my life are excellent.

_____ I am satisfied with my life.

_____ So far I have gotten the important things I want in life

_____ If I could live my life over, I would change almost nothing.

Here are a number of personality traits that may or may not apply to you. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

| Disagree Strongly | Disagree Moderately | Disagree a little | Neither Agree nor disagree | Agree a little | Agree Moderately | Agree strongly |
|----------------------|------------------------|----------------------|-------------------------------|-------------------|---------------------|-------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

I see myself as:

1. _____ Extraverted, enthusiastic.
2. _____ Critical, quarrelsome.
3. _____ Dependable, self-disciplined.
4. _____ Anxious, easily upset.
5. _____ Open to new experiences, complex.
6. _____ Reserved, quiet.
7. _____ Sympathetic, warm.
8. _____ Disorganized, careless.
9. _____ Calm, emotionally stable.
10. _____ Conventional, uncreative.

REFERENCES

- Adams, F. (1986). Intention and intentional action: The simple view. *Mind and Language*, 1, 281-301.
- Adams, F., & Steadman, A. (2004a). Intentional action in ordinary language: Core concept or pragmatic understanding. *Analysis*, 74, 173-181.
- Adams, F., & Steadman, A. (2004b). Intentional actions and moral considerations: Still pragmatic. *Analysis*, 74, 264-267.
- Audi, R. (2005). *The Good and the Right: A Theory of Intuition and Intrinsic Value*. Princeton: Princeton University Press.
- Bealer, G. (1992). The incoherence of empiricism. *The Aristotelian Society Supplementary Volume*, 66, 99-138.
- Bealer, G. (1996). A priori knowledge and the scope of philosophy. *Philosophical Studies*, 81, 121-142.
- Bealer, G. (1998). Intuition and the autonomy of philosophy. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 201-239). Lanham: Rowman and Littlefield.
- Berofsky, B. (2002). Ifs, cans, and free will: The issues. In R. Kane (Ed.), *The Oxford Handbook of Free Will* (pp. 181-201). New York: Oxford University Press.
- Bishop, M., & Trout, J. D. (2005). *Epistemology and the Psychology of Human Judgment*. New York: Oxford University Press.
- Bonjour, L. (1998). *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge: Cambridge University Press.
- Campbell, C. A. (1951). Is 'freewill' a pseudo-problem? *Mind*, 60, 441-465.
- Chamorro-Premuzic, T., Furnham, A., & Ackerman, P. (2006). Ability and personality correlates of general knowledge. *Personality and Individual Differences*, 41, 419-429.
- Claxton, G. (1998). Investigating human intuition: Knowing without knowing why. *The Psychologist*, 11, 217-220.
- Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory. *Psychonomic Bulletin and Review*, 13, 991-997.
- Cushman, F., & Mele, A. (in press). Intentional action: Two-and-a-half folk concepts? In J. Knobe & S. Nichols (Eds.), *Experimental Philosophy*. Oxford: Oxford University Press.

- Dennett, D. (1984). I could not have done otherwise: So what? *Journal of Philosophy*, 81, 553-565.
- DePaul, M. (1998). Why bother with reflective equilibrium. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 293-309). Lanham: Rowman and Littlefield.
- Descartes, R. (1985). *The Philosophical Writings of Descartes* (J. Cottingham, R. Stoothoff & D. Murdoch, Trans. Vol. 1). Cambridge: Cambridge University Press.
- Dollinger, S., & LaMartina, A. (1998). A note on moral reasoning and the five-factor model. *Journal of Social Behavior and Personality*, 13, 349-358.
- Ekstrom, L. (2002). Libertarianism and Frankfurt-style cases. In R. Kane (Ed.), *The Oxford Handbook of Free Will* (pp. 309-322). New York: Oxford University Press.
- Ericsson, A., & Simon, H. (1993). *Protocol Analysis: Verbal Reports as Data* (Revised ed.). Cambridge: MIT Press.
- Feltz, A., & Bishop, M. (in press). The role of intuition in naturalized epistemology. In M. Milkowski & K. Talmont-Kaminski (Eds.), *Beyond Description: Normativity in Naturalised Philosophy*. London: College Publications.
- Feltz, A., & Cokely, E. (2007). An Anomaly in Intentional Action Ascriptions: More Evidence of Folk Diversity. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 1748). Austin, TX: Cognitive Science Society.
- Feltz, A., & Cokely, E. T. (in press). The fragmented folk: More evidence of stable individual differences in moral judgments and folk intuitions. *Proceedings of the 30th Annual Cognitive Science Society*.
- Feltz, A., Cokely, E. T., & Nadelhoffer, T. (in press). Natural compatibilism v. natural incompatibilism. *Mind & Language*.
- Feltz, A., & Zarpentine, C. (manuscript). Do you know more when it matters less?
- Forsyth, D. (1980). A taxonomy of ethical ideologies. *Journal of Personality and Social Psychology*, 39, 175-184.
- Forsyth, D. (1981). Moral judgment: The influence of ethical ideology. *Personality and Social Psychology Bulletin*, 7, 218/223.
- Forsyth, D. (1985). Individual differences in information integration during moral judgment. *Journal of Personality and Social Psychology*, 49, 264-272.
- Forsyth, D. (1992). Judging the morality of business practices: The influence of personal moral philosophies. *Journal of Business Ethics*, 11, 461-470.

- Forsyth, D., & Berger, R. (1982). The effects of ethical ideology on moral behavior. *The Journal of Social Psychology, 117*, 53-56.
- Forsyth, D., Nye, J., & Kelley, K. (1988). Idealism, relativism, and the ethic of caring. *The Journal of Psychology, 122*, 243-248.
- Forsyth, D., & Pope, W. (1984). Ethical ideology and judgments of social psychological research: Multidimensional analysis. *Journal of Personality and Social Psychology, 46*, 1365-1375.
- Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19*, 25-41.
- Gettier, E. (2000). Is justified true belief knowledge? In E. S. J. Kim (Ed.), *Epistemology: An Anthology* (pp. 58-59). Malden: Blackwell.
- Gilovich, T., & Griffin, D. (2002). Heuristics and biases: Then and now. In D. G. T. Gilovich, & D. Kahneman (Eds.), *Heuristics and Biases: the Psychology of Human Judgment* (pp. 1-18). New York: Cambridge University Press.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge: Harvard University Press.
- Goldman, A. (2000). What is justified belief? In E. Sosa & J. Kim (Eds.), *Epistemology: An Anthology* (pp. 340-353). Malden: Blackwell.
- Goldman, A., & Pust, J. (1998). Philosophical theory and intuitional evidence. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 179-197). Lanham: Rowman and Littlefield.
- Goodman, N. (1955). *Fact, Fiction, and Forecast*. New York: Bobbs-Merrill Company.
- Hyde, J. (1990). Meta-analysis and the psychology of gender differences. *Signs: Journal of Women in Culture and Society, 16*, 55-73.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*, 341-350.
- Kane, R. (1999). Responsibility, luck, and chance: reflections on free will and indeterminism. *Journal of Philosophy, 96*, 217-240.
- Kaplan, M. (1994). Epistemology denatured. *Midwest Studies in Philosophy, 19*, 350-365.
- Knobe, J. (2003a). Intentional action and side-effects in ordinary language. *Analysis, 63*, 190-193.
- Knobe, J. (2003b). Intentional action in folk psychology: An experimental investigation.

- Philosophical Psychology*, 16, 309-323.
- Knobe, J. (2004a). Intention, intentional action and moral considerations. *Analysis*, 64, 181-187.
- Knobe, J. (2004b). Folk psychology and folk morality: Response to critics. *Journal of Theoretical and Philosophical Psychology*, 24, 270-279.
- Knobe, J. (2006). The concept of intentional action: A case study in uses of folk psychology. *Philosophical Studies*, 130, 203-231.
- Knobe, J., & Burra, A. (2006). The folk concept of intention and intentional action: A cross-cultural study. *Journal of Culture and Cognition*, 6, 113-132.
- Knobe, J., & Mendlow, G. (2004). The good, the bad, and the blameworthy: Understanding the role of evaluative reasoning in folk psychology. *Journal of Theoretical and Philosophical Psychology*, 24, 252-258.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-development approach to socialization. In D. Goslin (Ed.), *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.
- Kornblith, H. (2002). *Knowledge and Its Place in Nature*. Oxford: Oxford University Press.
- Leslie, A., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science*, 17, 421-427.
- Levin, I., & Gaeth, G. (1988). How consumers are affected by the framing of attribute information before and after consuming the product. *The Journal of Consumer Research*, 15, 374-378.
- Lewis, D. (1983). *Philosophical Papers*. Oxford: Oxford University Press.
- Malle, B. (2006). Intentionality, morality, and their relationship in human judgment. *Journal of Cognition and Culture*, 6, 87-112.
- Malle, B., & Nelson, S. (2003). Judging *Mens Rea*: The tension between folk concepts and legal concepts of intentionality. *Behavioral Sciences and the Law*, 21, 563-580.
- McCann, H. (1998). *The Works of Agency: On Human Action, Will, and Freedom*. Cornell: Cornell University Press.
- McCann, H. (2005). Intentional action and intending: Recent empirical studies. *Philosophical Psychology*, 18, 737-748.
- Mele, A. (1992). Recent work on intentional action. *American Philosophical Quarterly*, 29, 199-217.

- Mele, A. (2001). Acting intentionally: Probing folk intuitions. In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality* (pp. 27-43). Cambridge: MIT Press.
- Mele, A. (2006). *Free Will and Luck*. New York: Oxford University Press.
- Mele, A., & Cushman, F. (2007). Intentional action, folk judgments, and stories: Sorting things out. *Midwest Studies in Philosophy*, 31, 184-201.
- Nadelhoffer, T. (2004a). Blame, badness, and intentional action: A reply to Knobe and Mendlow. *Journal of Theoretical and Philosophical Psychology*, 24, 259-269.
- Nadelhoffer, T. (2004b). The Butler problem revisited. *Analysis*, 64, 277-284.
- Nadelhoffer, T. (2004c). On praise, side effects, and folk ascriptions of intentional action. *Journal of Theoretical and Philosophical Psychology*, 24, 196-213.
- Nadelhoffer, T. (2005). Skill, luck, control, and intentional action. *Philosophical Psychology*, 18, 341-352.
- Nadelhoffer, T. (2006a). On trying to save the simple view. *Mind and Language*, 21, 565-586.
- Nadelhoffer, T. (2006b). Desire, foresight, intentions, and intentional action: Probing folk intuitions. *Journal of Cognition and Culture*, 6, 133-157.
- Nadelhoffer, T. (2006c). Bad acts, blameworthy agents, and intentional actions: Some problems for juror impartiality. *Philosophical Explorations*, 9, 203-220.
- Nadelhoffer, T., & Feltz, A. (manuscript). The actor-observer bias and moral intuitions: Adding fuel to Sinnott-Armstrong's fire.
- Nadelhoffer, T., & Nahmias, E. (2007). The past and future of experimental philosophy. *Philosophical Explorations*, 10, 123-149.
- Nahmias, E. (2006). Folk fears about freedom and responsibility: Determinism vs. reductionism. *Journal of Cognition and Culture*, 6, 215-237.
- Nahmias, E., Coates, J., & Kvaran, T. (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions. *Midwest Studies in Philosophy*, 31, 214-242.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2004). The phenomenology of free will. *The Journal of Consciousness Studies*, 11, 162-179.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology*, 18, 561-584.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2006). Is incompatibilism intuitive?

- Philosophy and Phenomenological Research*, 73, 28-53.
- Nichols, S. (2002). How psychopaths threaten moral rationalism: Is it irrational to be amoral? *The Monist*, 85, 285-303.
- Nichols, S. (2004a). *Sentimental Rules*. New York: Oxford University Press.
- Nichols, S. (2004b). After objectivity: An empirical study of moral judgment. *Philosophical Psychology*, 17, 5-28.
- Nichols, S., & Knobe, J. (in press). Moral responsibility and determinism: The cognitive science of folk intuition. *Nous*.
- Nichols, S., & Ulatowski, J. (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind & Language*, 22, 346-365.
- Nisbett, R., Kaiping, P., Incheol, C., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, 108, 291-310.
- Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 64, 231-259.
- O'Connor, T. (2000). *Persons and Causes: The Metaphysics of Free will*. New York: Oxford University Press.
- Pereboom, D. (2001). *Living Without Free Will*. Cambridge: Cambridge University Press.
- Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology*, 17, 145-171.
- Phelan, M., & Sarkissian, H. (in press). The folk strike back: Or, why you didn't do it intentionally, though it was bad and you knew it. *Philosophical Studies*.
- Pink, T. (2004). *Free Will: A Very Short Introduction*. New York: Oxford University Press.
- Reid, T. (1997). *An Inquiry into the Human Mind on the Principles of Common Sense*. University Park: Pennsylvania State University Press.
- Reid, T. (2002). *Essays on the Intellectual Powers of Man*. University Park: Pennsylvania State University Press.
- Rest, J. (1979). *Development in Judging Moral Issues*. Minneapolis: University of Minnesota.
- Rusting, C. (1999). Interactive effects of personality and mood on emotion-congruent memory and judgment. *Journal of Personality and Social Psychology*, 77, 1073-1086.
- Shafer-Landau, R. (2003). *Moral Realism: A Defense*. New York: Oxford University Press.

- Siegel, H. (1984). Empirical psychology, naturalized epistemology, and first philosophy. *Philosophy of Science*, 51, 667-676.
- Sinnott-Armstrong, W. (in press). Framing moral intuitions.
- Smart, J. J. C., & Williams, B. (1973). *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.
- Smilansky, S. (2002). Free will, fundamental dualism, and the centrality of illusion. In R. Kane (Ed.), *The Oxford Handbook of Free Will* (pp. 489-505). New York: Oxford University Press.
- Smilansky, S. (2003). Compatibilism: The argument from shallowness. *Philosophical Studies*, 115, 257-282.
- Smith, M. (1994). *The Moral Problem*. Malden: Blackwell.
- Sosa, E. (1998). Minimal intuition. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 257-269). Lanham: Rowman and Littlefield.
- Sosa, E. (2007). Intuitions: Their nature and epistemic efficacy. *Grazer Philosophische Studien*, 74, 51-67.
- Sosa, E. (in press). A defense of intuitions. In M. Bishop & D. Murphy (Eds.), *Stich and His Critics*. Malden: Blackwell.
- Stanley, J. (2005). *Knowledge and Practical Interests*. New York: Oxford University Press.
- Stich, S. (1990). *The Fragmentation of Reason*. Cambridge: MIT Press.
- Stich, S. (1993). Will the concepts of folk psychology find a place in cognitive science. In S. Christensen & D. Turner (Eds.), *Folk Psychology and the Philosophy of Mind* (pp. 82-92). Hillsdale: Lawrence Erlbaum Associates.
- Stich, S. (1998). Reflective equilibrium, analytic epistemology, and the problem of cognitive diversity. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 95-112). Lanham: Rowman and Littlefield.
- Sverdlik, S. (2004). Intentionality and moral judgments in commonsense thought about action. *Journal of Theoretical and Philosophical Psychology*, 34, 224-236.
- Swain, S., Alexander, J., & Weinberg, J. (2008). The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research*, 76, 138-155.

- Szalma, J., & Hancock, P. (2005). Individual differences in information processing. In D. McBride & D. Schmorrow (Eds.), *Quantifying Human Information Processing* (pp. 177-193). Lanham: Lexington Books.
- Tidman, P. (1996). The justification of a priori intuition. *Philosophy and Phenomenological Research*, 56, 161-171.
- Tolhurst, W. (1990). On the epistemic value of moral experience. *Southern Journal of Philosophy*, 29, 67-87.
- Tolhurst, W. (1998). Seemings. *American Philosophical Quarterly*, 35, 293-302.
- Tversky, A., & Kahneman, D. (1982). Judgments under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovic & a. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases* (pp. 3-20). New York: Cambridge University Press.
- van Inwagen, P. (1997). Materialism and the psychological continuity account of personal identity. *Nous*, 31, 305-319.
- Weinberg, J., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29, 429-460.
- Williamson, T. (in press). *The Philosophy of Philosophy*.
- Wilson, M., Daly, M., Gordon, S., & Pratt, A. (1996). Sex differences in valuations of the environment? *Population and Environment*, 18, 143-159.
- Zelenski, J., & Larsen, R. (2002). Predicting the future: How affect-related personality traits influence likelihood judgments of future events. *Personality and Social Psychology Bulletin*, 28, 1000-1010.

BIOGRAPHICAL SKETCH

Adam Feltz received his Ph.D. in philosophy at Florida State University and his M.A. in philosophy at Northern Illinois University. His primary research interests are how folk intuitions are related to various philosophically important issues and how these folk intuitions can make practical differences in decision making. His articles have appeared in journals such as *Midwest Studies in Philosophy*, *Philosophical Explorations*, *Mind and Language*, and *Mind and Behavior*. He lives with his wife, Silke Feltz.