

INTUITIONS

Mark Kelman*

INTRODUCTION.....	1291
I. PROBLEMS IN THE SIMPLE STORY.....	1299
A. <i>How Significant Is the Resolution of the Underlying Philosophical Issue to Law?</i>	1299
B. <i>Why Might Particular Sorts of Commonsense Intuitions Matter?</i>	1307
1. <i>Diversity of intuitions</i>	1310
2. <i>Reflective intuitions</i>	1313
3. <i>Unreflective intuitions: moral competency</i>	1316
CONCLUSION.....	1322

INTRODUCTION

In *Commonsense Morality and the Ethics of Killing in War*, one of the many excellent papers presented at the Seventh Annual Conference on Empirical Legal Studies (CELS) here at Stanford, the authors set out to determine what it is that ordinary people actually think about whether one or another course of action that might be regulated by an international legal regime is actually the more moral course of action.¹ The paper's basic approach has been quite the rage over the past decade: experimental philosophers present vignettes to survey respondents and analyze how they respond to the vignettes. When surveying subjects about whether they think a particular action that legal regimes might permit, mandate, or proscribe is morally acceptable or desirable, authors typically advance two sorts of claims that the work they are doing is relevant to policymakers.

While it is the second of these kinds of claims that I will ultimately address in detail in this Essay, the first claim to relevance is critical to note as well, in part because this first claim is less controversial and in part because it is important to recognize how valuable this sort of empirical work really is even if

* James C. Gaither Professor and Vice Dean, Stanford Law School. Thanks to Dan Ho, both for the work he did organizing the CELS conference and for comments on an earlier draft.

1. Yitzhak Benbaji et al., *Commonsense Morality and the Ethics of Killing in War* (Apr. 8, 2013) (unpublished manuscript) [hereinafter *Killing in War*], available at <http://ssrn.com/abstract=2094022>.

only this claim to relevance were valid. The first claim could best be described as a functionalist claim about regime efficacy: it is important to know what commonplace moral beliefs are, because, all else equal, legal rules should track commonplace moral beliefs, even if these commonplace beliefs do not embody or even help us understand what is “truly” moral.² This is true in part because legal compliance and legitimacy flow from the intuitive plausibility or acceptability of substantive rules. People may more readily learn rules that they intuit would govern their conduct than they would learn counterintuitive rules; they may behave according to the rules they believe *should* govern behavior rather than those that actually *do*; they might not seek to enforce counterintuitive rules; they might well obey rules more readily even without regard to the expected selfish costs and benefits of obedience and disobedience if the rules we adopt comport with their moral intuitions; they might find a legal system more generally legitimate (and therefore respect even rules distinct from the precise rules at issue) if positive law typically matched beliefs. To the degree, too, that legislation ought to be responsive to popular will, vignette response gives us one sort of information about one relevant sort of public opinion. The authors of *Killing in War* indeed properly evoke these familiar functionalist justifications for caring about what commonplace morality demands.³

The second claim—that ascertaining what people’s moral intuitions are will help us determine what legal rules are truly normatively desirable—is the more contested one, and it is this contention that I will focus on in this Essay. There is a simple story—maybe, alas, a simplistic one—about the contributions that empiricists might make to the work that normative philosophers interested in law and lawyers interested in philosophy do. The simple story goes something like this: First, there are a substantial number of legal issues whose proper resolution is at least sensitive to, and perhaps even determined by, the resolution of some issue that would be thought of as “philosophical.” Second, many,

2. See, e.g., Jonathan Baron & Ilana Ritov, *Intuitions About Penalties and Compensation in the Context of Tort Law*, 7 J. RISK & UNCERTAINTY 17, 18-20 (1993); Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 454, 456-58, 471-78 (1997); Tess Wilkinson-Ryan & Jonathan Baron, *Moral Judgment and Moral Heuristics in Breach of Contract*, 6 J. EMPIRICAL LEGAL STUD. 405, 407 (2009).

3. As Benbaji and his coauthors state:

[I]t is preferable for legal convention to reflect the moral intuitions of individuals. Scholars distinguish between extrinsic compliance motivation, which is linked to actions driven by external commands, and intrinsic compliance motivation, which emerges from individuals’ own moral convictions. Intrinsically motivated individuals comply with the law even when the likelihood of punishment is low. When the law does not follow their moral intuitions, individuals comply only as long [as] they feel threatened by punishment.

Killing in War, *supra* note 1, at 23 (footnote omitted). There are a host of studies—of just the sort that characterizes the empirical legal studies movement generally—addressing whether these sorts of claims about obedience are *true*, rather than merely credible, but this literature is not my subject here. See, e.g., TOM R. TYLER, *WHY PEOPLE OBEY THE LAW* (1990) (examining what it means for citizens to find legal authority legitimate and investigating whether feelings of legitimacy increase compliance).

though by no means all, philosophers believe that the resolution of many philosophical issues depends on specifying and clarifying commonplace intuitions about how the issue ought to be resolved. Their preferred solution to a philosophical question is merely the articulated, principled answer that covers cases that ordinary people decide in a fashion consistent with the articulated principle, though the common folk frequently could not articulate the principle. Third, while armchair philosophers might believe they can intuit what commonplace intuitions really are—mistaking introspection for population sampling—they might well be wrong: experimental empiricists could therefore advance philosophical inquiry by helping philosophers and philosophically inclined legal academics both learn how widespread an intuition about a particular legally relevant philosophical proposition might be and explore how commonplace intuitions could best be specified, by subtly altering experimental prompts in relevant ways and observing the shifts in responses these reformulations elicit. The *Killing in War* paper is a good piece to use to explore some of this story's virtues and pitfalls. At the same time, I will make reference to some of the other fine papers at the conference, as well as other literature in the empirical philosophy tradition, in trying to illustrate some of the ways in which the basic story is incomplete, unclear, or unpersuasive.

Here is the basic argument in the *Killing in War* paper. International legal norms about *jus in bello*—rules about the conduct of war and more particularly rules regulating when it is permitted or forbidden to kill in war, rather than the *jus ad bellum* rules that govern when wars can be permissibly fought at all—do and should significantly derive from moral/ethical beliefs about the propriety of killing during armed conflicts.⁴ There is a normative philosophical dispute that the *Killing in War* authors believe can be understood as a debate between the “traditional” theorists” and “revisionists.”⁵ Traditionalists (most prominently Michael Walzer) believe that all killings of soldiers in wartime are justified, regardless of whether the soldiers are engaged in a just or unjust war, and that all intentional killings of civilians are unjustified (and that unintended side-effect

4. *Killing in War*, *supra* note 1, at 4, 8, 23-24.

5. *Id.* at 4-9. McMahan's “revisionist” argument is grounded in a number of propositions, both about the appropriate limits on the use of self defense against justified attackers, see Jeff McMahan, *The Ethics of Killing in War*, 114 ETHICS 693, 698-702 (2004), and the need that he sees to judge whether individuals are liable to harm by virtue of their own conduct, rather than to pay heed (as Walzer does) merely to one's membership in a protected or unprotected group. *Id.* at 694-96, 698-701, 733; Jeff McMahan, *The Morality of War and the Law of War*, in JUST AND UNJUST WARRIORS: THE MORAL AND LEGAL STATUS OF SOLDIERS 19, 19-22 (David Rodin & Henry Shue eds., 2008) [hereinafter McMahan, *Morality of War*]. Moreover, as Benbaji and his coauthors note, many writers on *jus in bello* take positions distinct from those they attribute to either side in the debate (for instance, that one cannot kill anyone, soldier or civilian, who does not pose a real, present threat to the person considering killing). See, e.g., Gabriella Blum, *The Dispensable Lives of Soldiers*, 2 J. LEGAL ANALYSIS 115, 115-16 (2010); Larry May, *Killing Naked Soldiers: Distinguishing Between Combatants and Noncombatants*, 19 ETHICS & INT'L AFF. 39, 39-40, 53 (2005).

killings of civilians should be minimized).⁶ This is true without regard to the underlying propriety of the war the soldier is fighting and is true even if a soldier is less involved rather than more involved in the war effort, or a civilian more rather than less involved. (The traditional view is sometimes labeled a contractarian position—despite the obvious objection that many or most soldiers are conscripted—because soldiers, like boxers, are participating in a social practice in which participants waive their ordinary immunities from even potentially fatal batteries.)⁷ The revisionist position, most associated with Jeff McMahan, is that the combatant/noncombatant line is not determinative: those people who are implicated in the prosecution of an unjust war should be vulnerable to the use of deadly force (whether soldiers or civilians highly enmeshed in the war effort) while those who are not (either because they are fighting with justification or because they are not truly implicated in the unjust war even if nominally serving as soldiers in the unjust belligerent nation's armed forces) should be immune from the use of force.

Both sides in the philosophical debate can be read to claim that the normative principles they are expressing are simply clear summaries or principled expressions of commonplace moral intuitions.⁸ So, the authors of *Killing in War* test whether, for at least one subpopulation (Israeli Jews), people in fact would, acting in the role of a military commander, order a lethal attack in circumstances that would (at least ostensibly, seem to) justify such an order under one principle but not the other.⁹ (They ask as well how they believe other Israelis would

6. See MICHAEL WALZER, *JUST AND UNJUST WARS: A MORAL ARGUMENT WITH HISTORICAL ILLUSTRATIONS* 36-37, 41-43, 136-159 (1977).

7. See, e.g., Thomas Hurka, *Liability and Just Cause*, 21 *ETHICS & INT'L AFF.* 199, 210-16 (2007).

8. The authors explicitly note that Walzer attacked the "revisionist" position by pointing out that it did not adequately account for people's intuitions that moral rules governing individual responsibility during wartime were distinct from those that governed in ordinary circumstances; in that sense, they imply that Walzer could be read to claim that it would undermine, if not fully rebut, the "revisionist" position to discover that it is counterintuitive. *Killing in War*, *supra* note 1, at 23. But the *Killing in War* authors are not as clear as would be ideal about whether they believe that either Walzer or McMahan themselves believe that it is important that the views they have propounded on these issues are intuitive, in some fashion or another. At the same time, the *Killing in War* authors do *not* themselves argue that a philosophical position is necessarily correct simply because it is intuitive. See *id.* at 23 ("[W]hatever the merits of the traditional and the revisionist views as moral theories, our findings suggest that commonsensical morality is sensitive to factors to which revisionists attach special importance."). Nor do they make the strong contrary assertion that intuitions are of little or no normative relevance.

9. *Id.* at 9-11. The authors dominantly employ a standard between-subjects design in which features of the vignette are manipulated; in this study, the authors manipulate the cause of the war (the subject acts as a commander making military decisions on behalf of an aggressor state, defender state, or a state fighting a war that arose from an ambiguous border dispute) and the categorization of the targeted individuals (combat soldiers, noncombat soldiers, involved civilians, and uninvolved civilians). *Id.* at 11-13.

Interestingly, the authors were unable in this experiment to present a vignette that evoked in study participants the archetype of a war in which neither side was clearly more to

act in the same circumstance, how they would evaluate the morality of the decision to engage, and how guilty they would feel if they had ordered an attack.) The gist of the findings is not especially important for my purposes, though some of the ways in which the findings are difficult to interpret may be. Nonetheless, it is worth reporting the authors' finding that while the targets' formal status as soldiers or nonsoldiers did matter a great deal (just as Walzer would predict), it was not exclusively determinative of responses (just as McMahan would surmise).¹⁰ Significant numbers of subjects also cared about the underlying justification of the war.¹¹ Finally, participants who believed that the war their side was fighting was unjust were prone to be especially scrupulous about following conventionally understood rules against harming civilians.¹²

The problems I will explore with the simple story can be summarized readily enough. The first proposition—that there are a substantial number of legal issues whose resolution significantly depends on the resolution of a philosophical quandary—may well be true, but it is equally important to recognize that the proper resolution of many legal issues is (and should be) either largely or wholly insensitive to the resolution of the philosophical debates even when the law deals with an issue that philosophers care about. As a first general approximation, I suspect that when people think that “philosophy” matters to law, they believe that a narrow branch of philosophy (defending normative propositions

blame than the other: they presented a conflict arising from a “long-lasting border dispute” hoping to evoke such a reaction, but a majority of the respondents thought one side was more at fault than the other in the dispute (presumably because Israeli Jews, the experimental subjects here, typically believe that the phrase “border dispute” is evoked by what they perceive as unjustly aggressive Palestinians). *Id.* at 12, 15. Moreover, slightly more than half the subjects failed to identify State B as just even when the experimental instructions noted that “State A, the stronger country, initiated war on State B, the weaker country, with the aim of taking over newly discovered gas deposits.” *Id.* at 11-12, 15 (internal quotation marks omitted). These findings suggest that it is often difficult to cabin real legal disputes in the terms philosophers use to describe them: there may be philosophical disputes about whether it is moral to kill those who are engaged in a just defensive war, but it is difficult for people to know *when* anyone in particular is engaged in such a war.

10. *Id.* at 15-17, 21-22. Note that it is far easier to *reject* the proposition that popular intuitions match Walzer's than it is to *verify* that they correspond with McMahan's. Since Walzer argues that nothing should matter other than the target's status as civilian or soldier, finding that other factors influence decisions for some participants falsifies his claim. But it is not nearly so clear that subjects account for additional facts in the ways that McMahan thinks they should (e.g., the degree to which a civilian is involved in the war effort or the degree to which the war is justified) in part because it is not clear what the apt degree of concern would be for anyone following his principles and in part because all that the data permits us to infer is that a subset of the population does indeed think other factors may be relevant.

11. *Id.* at 22, 33. In a prior draft of the piece presented at the Conference, the authors further found that subjects were concerned with the importance of the attack to the military campaign generally. See Yitzhak Benbaji et al., *Commonsense Morality and the Ethics of Killing in War* 18 (CELS Version, Nov. 2012) (on file with author) [hereinafter *Killing in War*, CELS Version].

12. *Killing in War*, *supra* note 1, at 18, 24-25.

about the morality of distinct personal actions) matters to law, and that legal rules should (with some limits) follow (or at least not violate) moral rules. But even setting aside the obvious point that philosophers are wholly uninterested in a wide range of legal disputes,¹³ it is important to remember that we may think it apt to resolve an issue for reasons skew to the reasons that concern philosophers, both outside the domain of rules grounded in moral norms and, for reasons I explore, even within the domain of morality-based rules.¹⁴

The second proposition—that intuitions, however intuitions are defined, count a great deal if our goal is to determine what is truly normatively desirable (at least, once more, when referring to the narrow subset of philosophical intuitions that might be thought of as norm-creating moral intuitions)¹⁵—is

13. Presumably, all legal rules must be justified (or criticized) within one or another broad philosophical framework (e.g., from the perspective of a utilitarian, a rule might seem inapt because adopting an alternative rule would increase aggregate social welfare), but the dispute over which of two rules would be more prone to increase social welfare might rely on empirical facts that do not bear on the general desirability of adopting utilitarianism as a framework so that qua philosophers, philosophers would not purport to contribute to the particular debate. So, for instance, philosophers may adduce general arguments for having (or not having) statutes of limitations for crimes, but would claim to add little to the debate over the optimal length of the statute. Those moral and political philosophers concerned with distributive ethics might have a particular view on how to trade off economic growth gains that redound to some with distributive gains that might redound to others, but would not add much to the debate about the impact of a particular marginal tax rate on incentives to hire.

14. I do not intend to rehash at length familiar debates over whether the law should, as a general matter, track morality or whether it is appropriate that our moral obligations be more extensive than our legal ones (e.g., demanding charitable outreach or good faith towards friends that the law might not). I do not ignore this point, but am more interested in pursuing some narrower points about the law/morality gap.

15. I suspect—though it is not really my main topic—that few people believe that there is any reason to believe that intuitions about a variety of jurisprudential issues matter in resolving these disputes (see, for example, the longstanding dispute about whether the source of law is merely the positive commands of a sovereign rather than some sort of natural law). Nor do I suspect that intuitions about a host of epistemological issues (some relevant to law, some not) have any privileged status in resolving these disputes. I see no evidence that anyone looks to see what lay people intuit about probability judgments when they argue about whether probability judgments about future events are best thought of as subjective (at core, just a series of bets one would accept or reject that one outcome rather than another will occur) or objective (at core, all future events are drawn from a preexisting sample of possible future events that resemble balls in an urn). Similarly, there are arguments about whether there are special reasons to believe factual propositions about the world, separate from the extent to which believing those propositions would advance one's other general subjective ends, and many epistemologists argue about these issues without making any reference to lay intuitions (or even "ordinary language") about the special features of "believing." See, e.g., Asbjørn Steglich-Petersen, *How to Be a Teleologist About Epistemic Reasons*, in REASONS FOR BELIEF 13 (Andrew Reisner & Asbjørn Steglich-Petersen eds., 2011). But there are a host of experiments designed to ascertain, for instance, whether subjects are more likely to believe that a character in a vignette knows something about a future state of the world—for example whether one *knows* a bank branch will be open this coming Saturday based on having seen it open in the past on Saturdays—depending on, for example, whether a mistake about the fact would be personally costly or whether the vignette highlights possible reasons

extremely controversial, and I will very briefly rehearse some of the arguments that we ought not care much what people intuit if we are trying to figure out what is morally desirable. (People may, for instance, simply intuit rules that happen to serve their narrow self-interest or match parochial, unstable prejudices.) I will focus instead on a distinct version of this second problem. What *sorts* of intuitions would most plausibly have normative weight and for what reasons should those sorts of intuitions count more than some other sort? Is an intuition about a moral problem distinct (and if so how) from a considered *belief* either about the proper resolution of the particular problem at stake or about what more general rules would best govern the resolution of the problem? Can we

that her inference may be undependable. For a good summary of these experiments, see JOSHUA ALEXANDER, *EXPERIMENTAL PHILOSOPHY: AN INTRODUCTION* 36-48 (2012). These experiments are never treated as illuminating epistemological truths, however.

What I think may be less immediately obvious but more important is that rather few people think lay intuitions about distributive justice issues are ultimately of much moment in resolving these issues, even though many works on distributive justice reason in part by offering examples that should appeal (intuitively?) to our sense of justice and even though intuitions about distributive justice could be said to be norm-creating moral intuitions. (For instance, libertarian theorists may argue that laypeople can be made to see the analogy between taxation and forced labor.) It is interesting to note that there were a number of papers at the CELS conference in which authors sought to describe popular sentiments on an issue that has plainly been addressed by those concerned with distributive justice, broadly construed, but do not imply in any way that popular responses to the problem have any particular normative status. Thus, for instance, Tamar Katz has found that experimental subjects intuitively accept wage penalties for members of protected groups to a significantly greater extent when the experimental manipulation leads the subjects to believe that the protected group members control the traits that often lead to discrimination in hiring or pay. Tamar Kricheli Katz, *Choice-Based Discrimination: Labor Force Type Discrimination Against Gay Men, the Obese and Mothers* 22 (CELS Version, Nov. 2012), available at <http://ssrn.com/abstract=2101596>. This can be thought to instantiate both a general (controversial) distributive principle—it is especially unjust if the distribution of resources proves sensitive to factors outside a person's control—and a particular (equally controversial) view about the nature of illicit labor market discrimination—people can justly suffer adverse consequences for possessing *irrelevant* traits towards which some employers bear irrational animus so long as they could avoid possessing the trait. But the author does not imply that the views of the experimental subjects are entitled to any normative deference (perhaps simply because the author does not share them?). Similarly Sanford Braver and his coauthors have plainly endorsed the public perceptions, whose existence they experimentally demonstrate, that child support grants should vary with the income of both noncustodial and custodial parents, but do not seem to believe that the existence of these intuitions about distributive justice within the family provide foundational support for the policy they prefer. Sanford L. Braver et al., *Public Intuitions About Fair Child Support Allocations: Converging Evidence for an "Ability to Contribute" Rule* 11 (CELS Version, Nov. 2012), available at <http://ssrn.com/abstract=2110376>.

For me, at this point, the most interesting question is why people at least sometimes do think some variety or other of intuition does matter in determining the proper answer to at least some set of ethical issues. Whether the explanation I proffer later in the text that certain sorts of intuitions might be thought of as especially significant in understanding how a moral issue ought to be resolved helps us explain why intuitions about distributive justice issues are generally not thought to determine the apt resolution of distributive justice disputes is a question I feel unable to resolve in anything but the most tentative way.

describe a person as having an intuition that it is right to do *X* if he has been taught that it is right to do *X*, and we believe that he could readily be taught that not-*X* is morally preferable, or must we believe that a belief is, at a minimum, rather recalcitrant before we class it as (the relevant sort of) intuition? Is a subject's "belief" about an apt legal rule (one sort of "general rule") still an intuition, of the appropriate sort, if she bases her belief in part on concerns about, say, how administrable one rule rather than another might be, even if she believes the less administrable rule is otherwise ethically or morally preferable to follow? Should we put *more* weight on intuitions if they are immune from reflection or put more stock in those that survive reflection? And on what basis might we say that unexamined intuitions are either adequate or even superior? It is this version of the second problem that is my single most pressing concern in this Essay.

This last set of problems leads me to question what it is that empiricists, relying above all on vignette-response surveys, actually tell us about the sorts of intuitions or beliefs they are reporting. If all forms of beliefs/intuitions were created equal, in terms of their relevance to judgments about the normatively apt resolution of those philosophical problems whose resolution mattered to law, then it hardly matters if we are learning about innate and immutable biological capacities to represent morally relevant problems in particular ways or learning about rules of thumb that subjects disclaim when made aware that they are using them or learning about the widely shared conventional responses that reflect learned local law and custom as much as they should provide a basis for them. But if only certain "sorts" of philosophical intuitions are normatively relevant, then we should be more careful in doing vignette-response surveys to identify which sort of intuition we are observing.

If I am right that only certain sorts of moral intuitions might even plausibly have normative force¹⁶—intuitions that are, in ways I come back to describe, grounded in unlearned basic competencies to represent moral problems in particular ways, relatively immune to reflection and revision—then the question is whether those doing vignette-response surveys should develop persuasive experimental techniques to help us label whether they have elicited such unlearned and unreflective responses rather than beliefs. I am certainly not professionally positioned to judge whether the existing techniques that might be used to distinguish what I am calling competencies from what I am calling beliefs are adequate to the task,¹⁷ and I am even less the one to figure out whether there might be better techniques.

16. I should note, up front, that I don't think even the most plausibly normatively relevant intuitions are actually of much normative moment, but I do believe that one can make a better case that one particular sort of moral intuition is relevant to figuring out what is and is not morally preferable than one can make the case that all sorts of intuitions, across the board, are relevant.

17. Psychologists interested in moral reasoning certainly do try to get at some of the issues I will argue are important—for instance, whether the subject can justify his use of an

I. PROBLEMS IN THE SIMPLE STORY

A. *How Significant Is the Resolution of the Underlying Philosophical Issue to Law?*

In the simplest version of the story that the sort of empiricism associated with experimental philosophy papers like *Killing in War* will transform normative legal discourse, we start with the supposition that the rule we choose should reflect our best understanding about what behavior is morally prohibited, mandated, or permitted. Thus, in this case, the supposition would be that international legal rules about when it is permissible to kill in the context of war should reflect moral truths about when killing is forbidden and permitted. (The further supposition, to which I return, is that moral truths are revealed by the exploration of some sorts of popular intuitions.) I will argue that it is not at all clear, generally, that legal rules should reflect the sorts of moral principles that philosophers study, and that, more particularly, it is not at all clear that rules about killing in war should mirror our philosophical beliefs about when killing is and is not morally permitted.

Now, of course, it is perfectly commonplace to recognize that there are many domains in which moral demands may be more stringent than legal demands, and for those predisposed to worry about an overreaching state, or to fret that inefficacious legal rules will crowd out and debilitate spontaneous private morality, those domains may be thought to be quite extensive. Nearly every first-year American law student is exposed to arguments that it might be inappropriate to hold people legally responsible, criminally or in tort, for failing to rescue the readily saved child drowning in a shallow pool of water, even if one believes that it is morally obligatory to rescue the child. (Of course, though, some of the arguments against legal liability in these cases are based on moral views that failing to save is *less* morally troublesome than actively harming another or that the legal system is morally bound to permit people to follow whatever life plan they wish so long as they do not directly violate others' rights to

intuitive category or not—merely by asking the subject whether he thinks he has good reasons for using the category or whether he subjectively feels confounded by why he reacts as he does. See, e.g., Fiery Cushman et al., *The Role of Conscious Reasoning and Intuition in Moral Judgment: Testing Three Principles of Harm*, 17 *PSYCHOL. SCI.* 1082 (2006); Marc Hauser et al., *A Dissociation Between Moral Judgments and Justifications*, 22 *MIND & LANGUAGE* 1 (2007). But I am completely agnostic as to whether seeking self-reports about the phenomenology of a vignette reaction is a reasonable way of ascertaining whether an intuition is of the relevant form. Similarly, developmental psychologists may well learn that certain morally relevant cognitions are developed before they could possibly have been learned or taught (and the sorts of intuitions that may prove most significant may be those that are understood before they could have been learned), but I am, once more, agnostic about whether developmental psychologists will be able to test whether cognate moral competencies undergird responses to each of the myriad of vignette responses that we have observed.

be immune from active harm.) It is quite hard to figure out when or whether anyone believes that there are moral duties as absolute and significant as the moral duties instantiated in legal rules that should be left outside the purview of law: I do find it quite plausible, though, that beliefs about the proper domains of private morals and public law are not based solely on judgments about the solemnity of the duty. Devout believers who disdain theocracy may well think that they have moral, rather than merely conventional, obligations to follow rules particular to their faith every bit as powerful as their obligation to comply with the moral rules that give rise to legal obligation, but still believe it would be a bad idea to impose these duties across the board in a pluralistic culture.

But the problem raised in interpreting the relevance of philosophy to law that we see in reflecting on the *Killing in War* piece is not much like the conventional problem that law is used to enforce only a portion of the moral code. I think the authors of *Killing in War* are right that there will, and should be, fairly complete legal rules governing the use of lethal force by commanders and soldiers, rather than some set of either separate or supplementary internalized, but fundamentally “private,” norms.¹⁸ Instead, the question of whether we can resolve legal quandaries once we have resolved the philosophical issues arises from three distinct, but interrelated, problems: First, the plausibility of the distinct general normative positions typically articulated as philosophical principles (for which distinct responses to vignettes are to serve as stand-ins) may depend far more than proponents of a particular position typically imply on suppositions about the nature of the cases likely to be covered by an abstract rule.¹⁹ Second, even when people have resolved broad normative issues, their inability to agree upon relevant facts may make legal issues remain controversial; it may well be that people’s normative disagreements frequently get

18. Nor is the problem that rules in this area have nothing to do with our moral norms. This is not, for instance, a dispute over an issue at the border between epistemology, cognitive psychology, and judicial administration: Can jurors adequately discount hearsay evidence in reaching factual judgments? Will jurors overestimate the prior probability of accidents that occur because of hindsight bias so that we need to take steps to avoid jurors finding defendants negligent that are not negligent given some set of rules about what negligence is?

19. In *Killing in War*, the “principles” the authors believe are competing are the “revisionist” and “traditional” positions. As I will note in the text, adherence to one or the other general position may be driven less by intuitions about the abstract content of each view than by reactions to exemplars that are never made explicit either by the experimenters or subjects. So the experimenters manipulate what they see as morally relevant categories—for example, am I dealing with a civilian or soldier? Is my potential victim fighting a just or unjust war?—and believe that those categories are doing work, with little regard to whether subjects exposed to the terms (I am fighting an unjust war) have the same sorts of conflicts or injustices in mind. A legal rule will have to cover concrete cases though, and without knowing more about how an abstract principle will cash out, it is not at all clear that we have settled legal controversies simply because we see some apparent consensus about a high-order principle.

displaced into disagreements about which factual claims are plausible.²⁰ Third, and most interesting in the case of the *Killing in War* paper, those who believe that legal rules are not self-executing may rightly pay considerably more heed to the possibility of implementing one rule, rather than another, even if they believe that the rule that is more readily implemented is not the one that reflects one's ideal moral sentiments.²¹

20. I understand that it is the aim of empirical legal studies *generally* to help resolve disputes over some of these contested facts. It is a more complicated issue than I can get into here when, whether, and to what degree empirical social science can increase consensus on ideologically charged and contested issues (e.g., the efficacy of the death penalty, efforts to diminish gun ownership in reducing homicide, or the impact of shifts in marginal tax rates on economic growth). My problem though is whether the particular, narrower sort of empiricism I am discussing in this Essay—the empirical study of popular philosophical intuitions—can help us resolve legal controversies.

21. I leave aside another significant variant of the problem that we may have reasons to believe that the answers to legal and philosophical problems *should* diverge: we may often believe that philosophers assume we have one set of ends when we may have another and thus believe their insights are more central than they are to our legal project, merely because their insights are unambiguously *topically related* to the legal issue at hand.

For instance, some may believe (I among them) that it is essentially unimportant how we resolve the question of whether criminals are morally responsible for their crimes in a (more-or-less) deterministic universe. Obviously, the question of the relevance of free will to blameworthiness has been a central preoccupation of both armchair and experimental philosophers (who have found that experimental subjects are far more prone to believe the compatibilist claim that people are morally responsible for misdeeds even in an imaginary world in which all conduct is absolutely determined, or at least is pictured as so predictably that some experimental subjects may treat it as determined, if the misdeeds are more morally horrific). See, e.g., Eddy Nahmias et al., *Is Incompatibilism Intuitive?*, in EXPERIMENTAL PHILOSOPHY 81, 86-89 (Joshua Knobe & Shaun Nichols eds., 2008). For summaries of a number of such studies, see ALEXANDER, *supra* note 15, at 16-17, 29-36. And these debates plainly *relate* to a set of familiar legal issues. But many would say that the criminal justice system can remain utterly agnostic on deep questions of whether blaming is “justified” (e.g., if it is simply operating as a social control system seeking, for instance, to incapacitate the dangerous). See, e.g., Lee S. Pershan, Note, *Selective Incapacitation and the Justifications for Imprisonment*, 12 N.Y.U. REV. L. & SOC. CHANGE 385, 390-91 (1983-1984) (favoring incapacitation of the dangerous rather than blame as the goal of the criminal justice system). Speculations about what the morally valid bases of blameworthiness are might be irrelevant even if the practice of *attributing* moralistic blame itself has “good” effects or if blaming some or all who cause harm is an unshakeable instinct. See, e.g., MICHAEL S. MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW 92-95 (1997) (discussing the positive effects of blame); Ernst Fahr & Simon Gächter, *Altruistic Punishment in Humans*, 415 NATURE 137, 137, 139 (2002) (analyzing whether blaming is intuitive or instinctual).

At the November CELS conference, there was a fascinating paper by Francis Shen that both traced the history of efforts to distinguish bodily and mental injury (in tort and criminal law) and explored, through experiments, the degree to which ordinary people neither find the distinction between mind and body clear nor resolve where to draw the line between bodily and mental injury in the same way. Francis X. Shen, *Mind, Body, and the Criminal Law* (CELS Version, Nov. 2012) (on file with author). Shen also proffers experimental evidence that potential jurors would be swayed by distinct jury instructions and the presentation of distinct forms of neuroscientific evidence to consider more conventionally “mental” injuries (like PTSD) as physical injuries. *Id.* at 51-52. My main point for now, though, is that while philosophers have plainly devoted a great deal of attention over the centuries to the

I will illustrate the first problem—the fact sensitivity of purportedly general conceptual propositions that philosophers may believe will dictate responses to the sorts of concrete problems lawyers must resolve—by brief reference to Judith Thomson’s well-known piece defending abortion rights, even on the supposition that the fetus is a rights-holding, living human being. Thomson asks us to imagine the following situation:

You wake up in the morning and find yourself back to back in bed with an unconscious violinist. A famous unconscious violinist. He has been found to have a fatal kidney ailment, and the Society of Music Lovers has canvassed all the available medical records and found that you alone have the right blood type to help. They have therefore kidnapped you, and last night the violinist’s circulatory system was plugged into yours, so that your kidneys can be used to extract poisons from his blood as well as your own. . . . To unplug you would be to kill him. But . . . it’s only for nine months. By then he will have recovered from his ailment, and can safely be unplugged from you.²²

She argues that a person would (plainly?) have the right to unhook the violinist, and argues that the fetus is just like the overly needy violinist and the pregnant woman seeking to terminate her pregnancy just like the kidnapping victim trying to get unplugged from the violinist. There are a host of familiar objections to the analogy (as well as counterarguments that the objections are normatively irrelevant or misleading) that are worth no more than a mention because they are mostly, but not completely, beside the point for my purposes here: one might, for instance, argue that abortion is more like killing and unhooking the violinist more like failing to save; that abortion is an intentional killing and the death of the unhooked violinist merely an unwanted but known side effect; that those who become pregnant as a result of voluntary intercourse rather than rape (sometimes conditional upon nonuse of inevitably imperfect contraception, sometimes not) are not involuntarily put in the position of someone needing to suffer in order to preserve another life; that the mother of a fetus has duties to save (as she would have duties to save her own drowning child) that she does not have towards strangers like the violinist.

But the point I want to emphasize—most relevant to the *Killing in War* paper—is that the force of the general principle that the case study is supposed to instantiate (that women should not be conscripted to save others, and that demanding that women bear the children they have jointly conceived is best seen as a form of illegitimate conscription) depends enormously on unstated factual suppositions about the burdens of both pregnancy and child rearing. I am not saying the suppositions are wrong—I, for one, would certainly agree that

mind/body problem, it is not at all clear that legal policymakers need care much (or even at all) about how the philosophical debate is resolved in thinking about what action constitutes a criminal assault (e.g., should verbal domestic abuse be considered an assault if we believe there is no significant mind/body distinction so that the suffering such abuse causes cannot be sharply distinguished, conceptually, from the suffering caused by striking the victim?).

22. Judith Jarvis Thomson, *A Defense of Abortion*, 1 PHIL. & PUB. AFF. 47, 48-49 (1971) (internal quotation mark omitted).

pregnancy, childbirth, and child rearing are indeed burdensome, so that absent strong countervailing desires to bear and/or raise a child, we would expect them to be strongly and justifiably resisted by many women. What I am saying is that it is not clear that the “general” principle one is supposed to glean from the argument—women’s bodies cannot be used against their will to sustain the life of a fetus, even if the fetus is deemed a living rights bearer—may depend on the factual supposition that the losses to the woman from the actual “use” of the body that pregnancy entails is significant rather than trivial and that it is not nearly so clear that everyone who endorses Thomson’s argument on abortion would endorse the more general version of the principle, as it might then be applied to the whole range of possible legal cases that could arise “under the principle” if it were thought to be generally determinative. (For instance, could a mother be forced against her will to give blood to sustain the life of a fetus she had jointly conceived that would be brought to term in a lab? The answer would appear to be “no” if the general principle adduced from the contemplation of particular facts were thought to be determinative across the range of cases, and the general principle would be thought to be that one has no duties to aid a fetus that one would not have to some random stranger, like the violinist.)²³

In the *Killing in War* paper, the notion that we genuinely employ distinct rules to cover cases involving soldiers versus civilians, or those fighting just, unjust, or ambiguous wars, or situations in which the decision to use lethal force is especially important to the war effort is problematic because our commitment to the general principle may well depend in part, in a quite similar way, on what we envision when we think of a soldier, or what we envision when we think of a just or unjust war. We see the problem explicitly when we reflect on certain aspects of the experiment—as I mentioned, respondents do not share the experimenters’ views about whether the particular states described

23. I suspect that the answer to *that* question, for many people, would depend somewhat on factual suppositions about the psychic consequences of knowing one has biological children, whether one is materially responsible for them or not. I suspect, too, that factual suppositions about what fetuses do or do not feel, with what sorts of consciousness, may impact views of what must be done to protect them, in ways that the abstract statement that Thomson concedes for argument’s sake (that they are rights-bearing living humans) does not adequately deal with. I believe, too, that many of the familiar controversies I adverted to about the Thomson article are also partly fact dependent rather than as abstractly normative as they are presented: for instance, the degree to which one treats pregnant women as “choosing” pregnancy depends not only on suppositions about the availability and efficacy of birth control but also on beliefs about the degree to which women are empowered in heterosexual relationships generally and these beliefs about the “degree of choice” may leak over and impact one’s beliefs about whether choice is a significant variable worth discussing. Similarly, it is difficult to know whether views about whether abortion is more like killing than letting die depend on particular views of the facts of how abortions are performed. Oddly, it would seem that later-term abortions could better be described as “letting die” cases because the fetus might not die until after it was removed from the womb, but the debates take place against a factual background in which the procedure itself is typically fatal.

in distinct experimental vignettes are fighting just or unjust wars. And we could see the problem more fully if the vignettes pressed further on the notion that we can readily identify a soldier, in ways all of us would agree to, especially in “postmodern” wars (conflicts?) fought only in part by explicit state actors employing a specialized group of uniformed soldiers. Many more people might assent to McMahan’s view that a soldier in a just war is immune from killing if they simply can scarcely imagine any situation in which their own country is the unjust one and the enemy’s soldiers are fighting the just war; their support for the abstracted version of the “rule” may mask enormous disagreement over what is ultimately the subject matter of legal regulation, the concrete cases it covers. And even as a general principle, it might not survive fuller exploration of the cases that the rule’s proponent intended to cover: they may seem to adhere to the principle only because they are not nearly adequately aware enough of the implications its proponents intend.

Even where experimenters try to eliminate this problem—as the authors do here by using vignettes that are meant to scream out that one country is in the wrong, and by placing subjects in a position where they are assigned no country towards which they owe long-term loyalty²⁴—they cannot readily do it as fully as would seem ideal. And it is troubling to suggest that they have shown us what normative position people actually hold on the underlying moral and legal questions in real cases when their actual positions are so sensitive to the scope and application of the rules they imagine.²⁵

It is simply much more difficult than conventional philosophers imagine to tell what “principles” people endorse, or to know if one has convinced them of

24. Obviously, the subjects are put in the hypothetical position of being commanders in one country’s armed forces and so that may be said to be their home country. And if one takes the social psychology literature on minimal group formation seriously enough, people identify with an arbitrary in-group pretty readily and quickly in experimental situations. *See, e.g.,* Samuel L. Gaertner & John F. Dovidio, *The Aversive Form of Racism*, in PREJUDICE, DISCRIMINATION, AND RACISM 61, 61-65, 67-73, 80-85 (John F. Dovidio & Samuel L. Gaertner eds., 1986). Still, it is difficult to tell not only whether these experimental subjects would behave the same way if Country *A* were actually named “Israel”—that problem is experimentally tractable, of course, since one could run these experiments with named countries—but I am quite doubtful that any lab setting creates the strong feeling of identification with one’s country’s position that people may typically develop when hostilities emerge, develop, and ripen into armed conflict between an “enemy” and one’s own country over a long period of time.

25. It might also be the case that beliefs supportive of Walzer’s view that all soldiers forfeit ordinary immunities from killing are in part dependent on unstated suppositions that soldiers volunteer (to a greater extent) for their positions, or that it is a long-established custom that young men, if not women, go through a period of vulnerability to death in war, or that civilian casualties will be avoided to a great extent if killings occur only on battlefields that are physically removed from the places that civilians live. So distinct ideas about facts may once more influence the acceptability of very broad norms, and yet we have little idea whether a person endorsing a general norm, or the application of what we see as the relevant norm to the vignette we present, is doing so because he has a particular exemplar in mind or view of how the world works.

the propriety of one's principles, when problems of application are so rampant that one can gain assent to a proposition merely because the ostensible assenter imagines a quite distinct set of practices will emerge if one follows the purported rule. I am dubious that the problem can be overcome readily, whether by armchair philosophers analyzing inexorably thin and decontextualized vignettes or, even more, by experimental philosophers who must present even thinner vignettes to subjects to get them to give tractable answers. Still, it plainly is, and should be, one of the long-term goals of experimental philosophers to elicit more and more information about what features of vignettes respondents actually care about and find salient in the vignettes to which they are exposed.

The second problem—that even people who truly agree on their norms or goals may have different views of how to achieve them—is familiar, whether one is considering goals with no particular moral content or heavily moralized goals. Each of us may agree that our sole goal is to get from Palo Alto to our dinner in San Francisco as fast as possible but disagree whether, right now, Highway 101 or Interstate 280 is the faster route. And each of us may agree that our only goal (in a particular policy domain) is to improve the long-term material welfare of our nation's poorest residents but disagree on which sorts of tax and income support programs will best achieve that goal. Even nonconsequentialists can be beset by such factual disagreements in their moral debates: starting from the premise that all those capable of exercising rational agency are justly punishable for violating laws, some might believe that psychopaths lack, and some may believe they possess, the relevant sort of capacity to engage in rational deliberation about their ends that make them subject to just punishment. (And if psychopaths lack the capacity, is it possible that others lack it if we look at their decisionmaking process in more detail?)

Once again, I think it is quite difficult to ascertain whether the disputes between the “revisionists” and “traditional” theorists highlighted in the *Killing in War* piece are simply disputes between people with different ideas about absolute decontextualized moral obligations or whether they are, at least in part, disputes between people who share (at least) some consequentialist goals (e.g., the reduction over time of the number of unjust wars, the resolution of conflicts with a minimum of death) but have unresolved and often unstated differences about whether one policy rather than another will get us there. And, as I noted, I am not sure that even the deontological accounts of why some killings are privileged and others impermissible are any more immune from being held hostage to the resolution of controversies over the apt description of the relevant characters: if, for instance, the—or at least a—critical issue for everyone is *whether* a person has forfeited his immunity from attack by agreeing to fight and be fought with, it is not at all clear that people will share views about whether all soldiers have made such agreements and whether all “involved” civilians have failed to (let alone whether non-uniformed, non-state-based “terrorists” have).

The third problem strikes me as the most interesting, perhaps generally, but certainly in the context of the *Killing in War* study. It strikes me as a perfectly plausible *legal* argument that a rule immunizing soldiers in just wars from being killed cannot be implemented in a world in which relatively few people believe their own country is being justly attacked.²⁶ (The question is *not* whether arguments about whether one rule or another is administrable are determinative, nor an argument that in this particular case, alternative rules—for example, rules that call on us to draw a clean distinction between soldiers and civilians—are any more administrable than a rule requiring us to determine whether a war is just.) And, of course, this study could be said to provide some evidence that people find it difficult to judge the justice of wars their own countries may fight in. (As I noted, Israeli Jews did not typically think that each state in a drawn-out border dispute were equally just, though the experimenters designed the vignette to elicit the response that they were. As I mentioned, it seems likely that the subjects believed instead that parties that don't disclaim longstanding border claims—presumably Palestinians in the minds of many Israeli Jewish subjects—are unjust.) But this suggests the possibility that we not only could believe that nearly all respondents intuitively believe that soldiers in just wars are immune from killing and believe that this intuition strictly defines our ideal moral duties but believe in the nonideal world of law, such an immunity rule could never govern behavior because judgments about the justice of the war would neither be readily shared nor authoritatively resolved in ways that would be acceptable to the parties who would be denied the legal authority to kill.²⁷

26. McMahan, of course, shares this concern and devotes a good deal of attention in his writing to distinguishing moral and legal rules of war. See, e.g., McMahan, *Morality of War*, *supra* note 5, at 27-43.

I refer to “implementation” problems as a quick capsule summary of a host of distinct problems. For instance, will those covered by the rule readily know when they are compliant and when they are noncompliant (and what are the costs of legal uncertainty)? Will distinct efforts to settle these sorts of disputes (e.g., by postwar trials of those determined to have killed soldiers fighting justly; by resolution before international bodies at the start of each armed conflict of the status of all combatants?) be seen as feasible or legitimate? Will these efforts cause new (unwanted) conflicts between nations?

The problem may be even worse than I imply in referring to the gap between “ideal” moral principles and administratively tractable legal ones. Assume that an act utilitarian may believe it is morally permissible to contract or exchange with anyone who will benefit (under some acceptable view of what it means to benefit) from the exchange, but that a rule utilitarian may believe it is *always* improper to exchange with people who are incompetent to contract (e.g., infants, the insane) even if one believes it is in the incompetent person's interest on the particular occasion to engage in the exchange, in part because the rule utilitarian worries that he, and others like him, will convince themselves too readily that the incompetent benefit from exchanges that actually benefit only the competent party. I don't think the competent party who follows an absolute unwavering moral rule that he will not contract with incompetent persons is following a “merely” legal, rather than moral, principle.

27. I return to the question of whether to treat a respondent who himself believes that there should be a gap between the legal rule and an ideal moral rule as having the relevant

B. *Why Might Particular Sorts of Commonsense Intuitions Matter?*

It is certainly possible to argue that as a matter of descriptive fact, philosophers do not actually care whether the views they advance are widely shared or “intuitive.” Certainly, many philosophers openly disclaim the idea that the force of the arguments they make depends in any way on whether the arguments match explicit intuitions about principles or best explain the intuitive responses to particular cases.²⁸ One suspects that many take delight in advancing positions that are counterintuitive.²⁹ What is more interesting, though, is to

sort of intuition about the case the experimenter poses. (What should we do with the respondent who cannot help but answer questions about what he would do or what he would feel is moral to do by reference to what he thinks he should be legally permitted to do and believes that the legal permissibility of conduct depends in significant part on whether a particular rule can be implemented without undue costs?) For now, I am assuming for exposition’s sake that the experimental subjects are giving a “purely” idealized moral answer and asking whether idealized moral answers determine appropriate legal outcomes.

28. The point is put nicely by Kwame Anthony Appiah, reflecting on intuitionism: For many, it is a point in Jeremy Bentham’s favor that, in contemplating the principle of utility, he got right the big issues that his contemporaries got wrong: he was able to challenge the prevailing moral intuitions of his day about slavery, the subjection of women, homosexuality, and so forth.

KWAME ANTHONY APPIAH, *EXPERIMENTS IN ETHICS* 77 (2008).

29. Just as it is the case, though, as I am about to explore briefly in the next note, that philosophers who *purport* to rely on intuitions may not *really* do so, it is also the case that those who purport to disdain them might actually rely on them at some point. It all depends, of course, on whether it is possible to claim that a moral argument is true or plausible without relying on the intuitive credibility or plausibility of some of the arguments on which the ultimate argument is grounded. So, for instance, it may be rather obviously true that some philosopher states it is morally obligatory to do *X* knowing full well that most people would not find that to be the case. But her argument that it is morally obligatory to do *X* may rely on why it follows from *Y* that *X* is obligatory and it may be the case that unless one shares the intuition that *Y* entails *X* being obligatory or that *Y* is itself true, the argument won’t be sensible. (Obviously, there may be more than two steps in the argument.) And it is possible as well that the veracity of some proposition that grounds a further proposition would remain opaque to most commonsensical listeners—it involves, for instance, fairly sophisticated mathematical understanding of logic that most people lack—but one might *still* believe that if the proposition could be unpacked for the ordinary commonsensical reader, he would understand it to be true.

So take for instance one of the most prominent nonintuitive propositions in late twentieth century moral philosophy, John Taurek’s claim that the numbers should not count in deciding whether to save one (larger) group of persons from death rather than a (smaller) group of persons when one does not have the time or resources to save both groups. Taurek argues that even if one were deciding how to allocate one’s lifesaving resources or capacities in a situation in which one believes it is not obligatory to save anyone at all, one should simply ignore how many people one is saving. Thus, assume that five people are drowning on your left and one person on your right, and you have time either to save the five or to save the one, but not to do both. If you had no ground specific to you as an agent on which to choose any person over any other person (e.g., one potential victim was your child), you should flip a coin to determine whether to save the one or the five since it would show equal respect for the sanctity of each wholly separate life if each person had the same fifty-fifty chance of surviving the ordeal. John M. Taurek, *Should the Numbers Count?*, 6 PHIL. & PUB. AFF. 293 (1977). In this view, showing a preference for saving five implies that we should be maxim-

explore whether philosophers who self-consciously claim that they care about intuitions actually do so. Interesting as this task is, though, it is not really the task I have set for myself.³⁰

izing the well-being of a fictitious collective aggregate. We could readily understand why a particular person would choose to sustain just one broken leg in her life rather than five. But there is no parallel way to say that it is better that one person break her leg than that five do because there is nobody from whose viewpoint this outcome is better, no being who experiences more broken bones than any other. *Id.* at 303-04, 307. Taurek's argument in this regard closely tracks the argument in ROBERT NOZICK, *ANARCHY, STATE, AND UTOPIA* 32-33 (1974).

I am certain that Taurek did not believe that this claim, put like that, was intuitively appealing. And in experimental work I have done with Tamar Kreps, we have empirically confirmed the predictable finding that far fewer people believe it permissible to save one drowning man rather than ten if one foregoes saving the ten by saving just the one. See Mark Kelman & Tamar Admati Kreps, *Playing with Trolleys (I): Intuitions About Aggregation*, 11 J. EMPIRICAL LEGAL STUD. (forthcoming 2014), available at <http://ssrn.com/abstract=2138475>. And I suspect, but am not certain, that Taurek did not believe that the most basic set of arguments that he offers to justify his claim would resonate with people's intuitions either. And my work with Kreps confirms the idea that even people exposed to the basic argument still believe it is far more permissible to save many than think it permissible to save one; in fact, exposure to the argument does not change subjects' responses to the vignettes at all. But that does not tell us whether he believes that the arguments he offers seem relevant, or troubling, only because they appeal to (at least some readers') intuitions about the foundational separateness of persons or the perils of thinking in terms of aggregated social welfare functions.

30. For an especially lucid exposition of the claim that philosophers do not in fact care about intuitions (even those philosophers who *state* that one's philosophical propositions should either directly match intuitions or, when applied rigorously to cases, match intuitive reactions to those cases), see HERMAN CAPPELEN, *PHILOSOPHY WITHOUT INTUITIONS* (2012). Obviously, his argument is more complex than this—it takes a book, not a footnote, to explicate—but one can get some of the flavor of the argument by considering why one might reject each of two distinct arguments that philosophers rely on intuitions. The first argument is that they seem to say that they do; they use the word “intuitively” to modify many of their propositions. But Cappelen argues that if one reads the pieces that use the word rigorously, one sees that the authors are not generally using it to mean that the argument is correct because it is intuitive. Rather, they most typically mean to hedge or disclaim the force of the argument they are about to offer. (So *X* may say, “Intuitively, *Y* is true,” as a way of saying that *Y* seems true as a first, not-carefully-considered observation.) *Id.* at 16-19. And in situations in which the word “intuitive” is not used as a hedge, it is used rather loosely and inconsistently. *Id.* at 52-60. He claims, at core, that philosophers have started casually stating that intuitions matter for some time now, but they don't really mean it nor does the statement affect their work.

Philosophers' use of “intuition” is a kind of intellectual/verbal virus (or tic[]) that started spreading about thirty to forty years ago. It is a bad habit and we should abandon it. However, and this is important, the virus didn't have much effect on first-order philosophy. . . . The most damage caused by the “intuition”-virus was on philosophical methodology. The virus helped convince those doing methodology that things called ‘intuitions’ play an important part in philosophical arguments.

Id. at 50.

The second argument is that they actually rely in their practice on intuitive reactions to the cases they offer when using a case method. But again, Cappelen demurs. So, for instance, take an example I mentioned earlier: Judith Thomson's use of the plugged-in violinist hypothetical in trying to recast the abortion debate from a debate focused exclusively on whether

Instead, what I want to explore is whether philosophers *should* care about intuitions, and, more importantly, why they should care about some sorts of things that could be described as intuitions more than others. My basic claim is that there is very little reason for normative philosophers to be especially interested in any of a long list of commonplace reactions that subjects might have to vignettes raising moral issues (even if, to track an earlier point I raised, lawmakers may care whether they are enacting laws grounded in widely shared beliefs, whether because they think such laws are more likely to be seen as legitimate and worthy of obedience or because the lawmaker has some commitment to democratic responsiveness). They should not care about reactions that depend heavily on what the subject has been taught; they should not care about reactions that vary a great deal across persons depending on culture³¹ or background beliefs.³² Arguably, they should not even care much about moral reactions that are the product of reflection and deliberation.³³ If they should care about anything, they should care about reactions—assuming they exist—that are most immune from reflection and revision, that are least understood by the subjects who have them, and that reflect basic shared human competencies. I am skeptical that it makes sense for normative philosophers to care about these

the fetus is a human being with the rights that humans more generally have. Cappelen claims that Thomson is not relying on the intuitive plausibility of the proposition that one could unplug oneself from the violinist or the intuitive plausibility of the analogy between the violinist case and the abortion case. Quite to the contrary, he notes that she states right from the start that her aim is merely to force a closer examination of an issue—an issue that will remain vexing—than snap (intuitive?) judgments would permit. *Id.* at 149-50. Her real goal is to highlight otherwise unobserved analytically relevant features of the problem, and the fact that reactions to the violinist case might have a particular sort of phenomenology or feeling does not work in her ultimate argument. *See id.* at 153.

31. This is not to say that it is not extraordinarily interesting to find out that reactions to a host of issues vary across cultures and to figure out what this tells us about how we develop beliefs. Learning that reactions vary cross-culturally may also serve a welcome cautionary role for those predisposed to trust their own reactions a bit too much. “[D]iversity in moral norms was an important catalyst to philosophical reflections about the status of our moral norms Where we do find diversity, . . . we can ask more informed questions about the relative merits of these different ways of thinking about the world.” Joshua Knobe & Shaun Nichols, *An Experimental Philosophy Manifesto*, in EXPERIMENTAL PHILOSOPHY, *supra* note 21, at 3, 11. We find diverse responses not only when we explore moral reactions but epistemological ones as well. For a good summary of some of the literature on epistemological diversity, see Jonathan M. Weinberg et al., *Normativity and Epistemic Intuitions*, in EXPERIMENTAL PHILOSOPHY, *supra* note 21, at 17, 23-34; and Edouard Machery et al., *Semantics, Cross-Cultural Style*, in EXPERIMENTAL PHILOSOPHY, *supra* note 21, at 47. My claim is the narrower one that we will not find much of *normative* interest when we observe such varied reactions.

32. I will note briefly some very tentative ideas about whether variation in reactions by gender ought to be seen as negating the relevance of “intuitions.” *See infra* note 58.

33. If one assumes that there is at least some significant class of cases in which reflective reactions differ from hasty, unreflective “blink” reactions, it might be the case that this provides an argument *against* attending to unreflective reactions in those cases, even if it does not provide an affirmative argument for attending to the reflective ones. I return to this point in the discussion of unreflective intuitions.

sorts of unreflective universal intuitions either—nor am I sure they exist, or that, to the degree they exist, they determine responses to actual moral quandaries—but they are the only sort of intuitions I can readily understand the case for attending to at all if one’s goal is to adopt the intuitions of others as one’s own preferred moral beliefs.

1. *Diversity of intuitions*

The most common argument against accounting for “local” intuitions is that they are prone to reflect ephemeral misconceptions, or the self-justifications of those seeking to bolster their own power, privilege, or self-interest.³⁴ But even if one did not worry, say, that “intuitions” about the appropriate status of the disadvantaged were not just apologia for the local brand of oppression, one might still be puzzled about why any non-universal intuition ought to have normative weight.

The basic argument against giving particular normative weight to reactions that are non-universal because they are learned or vary across cultures is that such reactions can best be understood as the outputs of just the sorts of debates that philosophers have over norms, rather than as inputs into the debates. If, for instance, everyone taught by *X* believes *Q* and everyone taught by *Y* believes *R*,³⁵ all we have really done is discover that people can be persuaded by the substantive arguments *X* offers or those that *Y* offers (or that people believe what they are told, whether “persuaded” or not). But even if we could learn something normatively important by finding that people *resist* learning certain propositions—more on that later—it is hard to see what we could learn about the normative force of one of two arguments, each of which we observe that people can learn. And the fact that, say, response *Q* were more common than response *R* might merely reflect that more people were exposed to “teacher” *X* than “teacher” *Y*.

Presumably, in cases in which responses vary because of exposure to explicit education or pervasive cultural practices, one set of “norm-educators” has come to believe one set of propositions and another believes contrasting ones. But the commitment to these propositions cannot be grounded in intuitions that precede their acquisition of the belief if widely held intuitions, or at least the intuitions that we now observe when asking people to respond to vignettes, don’t get firmly established until someone comes along to educate people about what to believe about the issue. So it would seem that the “norm-educators” came to their beliefs in the same way that we could now, without regard to vi-

34. Again, Appiah makes the point economically: “The suspicion that our common sense may be littered with perishable and parochial prejudice is . . . an ancient and enduring one . . .” APPIAH, *supra* note 28, at 81.

35. We can think of a teacher in this regard as a particular person or as a wide range of acculturating institutions.

gnette responses, and there is no particular reason to accept their resolutions simply because the beliefs they had have been at least reasonably widely disseminated while the contrary ones we prefer have not yet been widely taught. This argument holds too, in much the same way, if what are meant to be bedrock intuitions vary across persons depending on conventional diversity of ideology or background beliefs.³⁶

The question of whether vignette responses should be thought of as learned (explicitly or by exposure to shared acculturating forces) is raised quite sharply in the version of the *Killing in War* paper presented at CELS. Those respondents who served as combat soldiers in the Israeli Army were significantly more likely to object to an attack on uninvolved civilians than an attack on any other target, those who had served as noncombat soldiers were significantly more likely to object to an attack on involved civilians than an attack on combat or noncombat soldiers, and those who had not served in the military at all were not significantly affected by the categorization of targets as involved or uninvolved or civilians or soldiers.³⁷ While it is not unambiguously the case that military veterans respond differently because they have been explicitly taught the conventional rules of war—in which the categorization of targets is tremendously important—it is certainly plausible that much of what we are learning about subjects' responses essentially reflects what they have been taught. It would be odd to say that we could justify existing laws of war as reflecting the moral “intuitions” that had simply been acquired by exposure to the laws we sought to justify. If I once believed *X* and convinced you that *X* is true, it would be peculiar to think that the reason I now believe that *X* is true is that you do.³⁸

36. The authors do not test for (nor find) such ideological diversity in the *Killing in War* piece, but if one imagines, say, that responses to the vignettes varied significantly depending upon party affiliation, one would, once more, think that distinct respondents with each reaction to a vignette had come to his or her view through the same sort of all-things-considered legal and policy judgment as one would come to on one's own, without making any reference to intuitions.

The same argument would almost surely apply if respondents' judgments were grounded in views over how readily one could implement one legal norm rather than another. See *supra* note 27. There is no reason that a show of hands on the best outcome of that sort of policy analysis would be entitled to any moral deference merely by virtue of the fact that one view was typically preferred to another.

I suspect that one of the main reasons that researchers do not typically imply that commonplace judgments on distributive ethics issues are entitled to normative deference, see *supra* note 15, is that such judgments are thought to vary strongly with ideology.

37. *Killing in War*, CELS Version, *supra* note 11, at 23.

38. The possibility that existing legal rules may determine moral reactions to vignettes, rather than that the reactions precede the formation of law and can serve as a basis for law, is raised not just in the *Killing in War* piece but in another fine piece presented at the November conference. In *Law, Norms, and the Motherhood/Caretaker Penalty*, Catherine Albiston and her coauthors find that when the existence of the Family Medical Leave Act was made salient to experimental subjects, they neither penalized mothers and caretakers (in wage terms or in terms of promotions) nor made negative normative evaluations of them as workers when workers in vignettes were either parents or leave takers rather than nonparents.

Since the *Killing in War* paper does not compare subjects across cultures, it is impossible to say whether responses to the moral issue are culture specific. But certain responses we do observe—most particularly, the subjects’ resistance to treating a “border dispute” vignette intended to portray a war in which neither side was more justified than the other as a neutral war, perhaps because Israeli Jews think about their conflicts with Palestinians when asked to consider a longstanding dispute over ownership of land—suggests that particularized local experiences may affect vignette responses. One could certainly imagine that respondents within a culture that had been confronted with a good deal of “terrorist” conflict in which nonsoldiers not especially associated with particular states had inflicted a good deal of violence might be less prone to draw sharp soldier/civilian distinctions in deciding whom one might justly kill.

But this sort of cultural variation (might one describe it as “learning from circumstances”?) is not precisely like the sort of cultural variation that I just described, which I would classify as teacher-responsive variation (learning from diffuse sources within the culture what is proper to believe). It is no longer the case that we can apply the simplest argument against accounting for learned intuitions—that one is merely indirectly accounting for some other party’s non-intuition-based normative beliefs that are entitled to no more presumptive deference than one’s own beliefs. Instead, the argument against accounting for these learned-from-circumstances intuitions takes the following form: it may be that if we observe culturally distinct intuitive reactions to a particular vignette, we need to reformulate our view of what the basic intuitive primitives really are. They may better be understood not so much as reactions to the vignette, but as tendencies to react to vignettes given certain experiences which in this case happen to be shared. (So imagine, quite counterfactually, that we observed that everyone in culture *A* were allergic to wheat and everyone in culture *B* to rice; it might be the case that the right generalization is that everyone in all cultures is allergic to whichever food they were exposed to first and that those in *A* first encountered wheat. There would be a basic biological trait—parallel to a basic universal intuition—but it is not the particular allergy—parallel to a concrete response to a vignette—but the process of developing the allergy that is the basic trait.)

Catherine Albiston et al., *Law, Norms, and the Motherhood/Caretaker Penalty* (CELS Version, Nov. 2012). On the other hand, subjects in their experiments for whom the existence of voluntary organizational policies were made salient were sometimes unaffected and sometimes made more likely to make discriminatory judgments, and subjects in prior experiments they advert to often penalized mothers and caretakers when neither law nor voluntary organizational policies were highlighted. And experimental subjects’ beliefs about philosophical propositions can be taught by nonlawyers too: Shen demonstrates in *Mind, Body, and the Criminal Law* that subjects’ views of the distinction between mental and bodily injury change when exposed to arguments by neuroscientists that call conventional distinctions into question. Shen, *supra* note 21, at 51-55.

2. *Reflective intuitions*

That leaves open the question of whether one should be more, not less, prone to credit reflective reactions than unreflective ones, on the supposition that these sometimes differ. I want first to note some of the basic arguments on each side of that controversy, but return to the issue when I discuss what I think the best arguments might be for paying special attention to unreflective gut reactions.³⁹

The commonplace arguments that we should be uninterested in experimental subjects' reflective responses, or at least not especially interested in them are: First, if we really care what people think after they have considered all sides of an issue and carefully tested whether their initial beliefs persist once they have considered the range of judgments that they would feel compelled to make across a wide range of cases if they held fast to their initial position, we might as well ask people ("experts," "philosophers") who are especially good at doing what we are now claiming is the relevant sort of self-critical reflection.⁴⁰ Second, it may well be that we simply think differently about concrete and abstract problems; each sort of response is intuitive *within its sphere*. A problem posed in a more abstract way may elicit our abstract reflection-based responses—that is to say, may activate brain regions or structures or distributed processes more associated with higher order cognition or reflection—but the reflective response does not elicit a more moral reaction, rather than the reaction we may bring to the table in our rule-giving, abstract mode. There are reasons to believe that our rule-giving mode is superior, reasons to believe it is inferior, and reasons to believe that it is just distinct.⁴¹

39. I leave aside the possibility that we might best *define* an intuition as a reaction with a particular phenomenology inconsistent with the idea there can even be a "reflective" intuition: some might say an intuition *is* a snap judgment, arrived at effortlessly, grounded in a feeling rather than a considered belief, and perceived as leading to *necessary* conclusions rather than conclusions subject to revision or critique. For discussions of some of these views of what intuitions are, see, for example, CAPELLEN, *supra* note 30, at 31-35, 65-66; and ALEXANDER *supra* note 15, at 21-23. I also set aside the precisely contrasting possibility that "raw intuitions"—yes/no reactions to philosophical propositions that accompany vignettes—are not truly *philosophical* intuitions because philosophical intuitions are best understood as referring only to intuitions about the veracity of abstract propositions, rather than particular controversies, *see id.* at 24, or as the subset of intuitions upon which we reflect critically, taking effort to insure that our judgments are not influenced by irrelevant considerations, *see id.* at 25-26.

40. For standard discussions of what is usually dubbed the "expertise" critique of the idea that philosophers ought to care about laypeople's intuitions, see ALEXANDER, *supra* note 15, at 90-98; and Knobe & Nichols, *supra* note 21, at 8-10. The expertise defense might be seen to imply—falsely in some views—that *philosophers* but not ordinary folk can rely on their more practiced intuitions, rather than that the methods philosophers use to explore propositions do not ultimately rely on intuitions at all. This latter point is emphasized in CAPELLEN, *supra* note 30, at 227-28.

41. *See, e.g.*, Walter Sinnott-Armstrong, *Abstract + Concrete = Paradox*, in EXPERIMENTAL PHILOSOPHY, *supra* note 21, at 209. Sinnott-Armstrong argues that when

There are also several commonplace arguments that we need to attend to reflective intuitions. First, even if there is an intuitive moral competence that would, in ideal circumstances, permit each of us to discern moral truths, we would expect performance errors to compromise our capacity to see those truths—just as we would expect performance errors to compromise our capacity to perceive, remember, or infer accurately. Thus, for instance, it should come as no shock that people respond to moral (and even epistemological) problems differently when they are more rather than less emotionally stirred by the problem since affect clearly affects performance on all mental tasks. We should examine reflective responses since it is these responses that best reveal our underlying competencies,⁴² and we should be as suspicious of first-line moral intuitions as we are of other immediate heuristics-based reactions, which may well serve us in good stead in the run of cases, but lead us to make errors of judgment when used across the board.⁴³ Second, philosophical and moral

problems are stated abstractly, we are more likely to be blame-denying determinists, strict consequentialists, and epistemological skeptics, but when looking at more concrete problems, we are more likely to blame, to follow deontological rules restraining us from taking at least some steps whose aggregate consequences are good, and to believe we know things even when we can imagine limits on our capacity to be certain of our claims to know. *Id.* at 216, 219, 221-22, 225. He suggests, rather tentatively, that our “intuitions” in concrete cases may play precisely the opposite role that they play for Cass Sunstein. *Cf.* Cass R. Sunstein, Lecture, *Moral Heuristics and Moral Framing*, 88 MINN. L. REV. 1556, 1556-58 (2004) (arguing that abstract reasoning is needed to override case-specific intuitions that are generally apt, but wrong in particular cases). Sinnott-Armstrong views it as plausible that the abstract responses are good heuristic approximations but that they must be narrowed in application by reaction to concrete cases where the general rule seems inapt. *See* Sinnott-Armstrong, *supra*, at 224. But he also seemingly views it as plausible that *neither* reflective nor unreflective reactions have any particular claim to validity, but that our recognition that reactions differ to each sort of problem should simply help explain why we feel conflicted over so many moral issues. *Id.* at 225-26.

42. This is one interpretation of the argument in Ernest Sosa, *Experimental Philosophy and Philosophical Intuition*, in EXPERIMENTAL PHILOSOPHY, *supra* note 21, at 231, 237 (“One would think that the ways of preserving the epistemic importance of perception in the face of such effects on perceptual judgments would be analogously available for the preservation of the epistemic importance of intuition in the face of such effects on intuitive judgments.”). *See* also, more generally, Antti Kauppinen, *The Rise and Fall of Experimental Philosophy*, 10 PHIL. EXPLORATIONS 95, 96-97 (2007).

43. Once more, Appiah’s description of this problem is well put (and is critical in thinking about the mirror image claim to the one we are discussing here, to wit that unexamined intuitions are most normatively interesting):

Understanding where our intuitions come from can surely help us to think about which ones we should trust. Other psychological research will suggest that some of our intuitions will survive, even in circumstances where they have misled us. Here the analogy with the scientific study of our cognitive processes is . . . natural. . . [A]n awareness of optical illusions helps us avoid being misled by them The proposal that . . . it’s our feelings that guide us to the intuition about the footbridge case, while our reason guides us in the original trolley problem is the sort of thing we might want to consider in deciding whether that intuition is right.

APPIAH, *supra* note 28, at 110-11. For an extensive discussion of the possibility that we employ rather recalcitrant moral heuristics that lead to error if used across the board, *see*, for

argument is not based solely on assertion; it is based on giving the sorts of reasons that only a reflective actor could provide. This is true not just because *public* discourse rightly demands reasons, but because even when we are trying to figure out what sorts of persons we aspire to be, we typically try to give reasons to embrace certain self-conceptions.⁴⁴ More generally, there are a host of reasons—summarized well by John Rawls in arguing that our most trustworthy beliefs are those that emerge in “reflective equilibrium” rather than those that reflect unprocessed intuitions⁴⁵—to believe that the views we would be most prone to characterize as moral are those that we come to only under certain conditions (we may try to think about rules we would favor if deprived of knowledge of our own circumstances or particular self-interest; we may try to think about rules that we would favor having confronted the best arguments for alternative rules).

I am ultimately agnostic about the question of whether reflective beliefs are indeed worthy of more deference than nonreflective ones, but I am unambiguously skeptical of the claim that we would rely on vignette-response intuitions that attempted to push subjects towards greater levels of reflection, rather than what is generally referred to as the reflective reactions of philosopher/experts, if we were trying to identify situations in which intuitions were indeed reflective enough, or reflective in the right way. Obviously, the *Killing in War* authors do not even attempt to force any sort of reconsideration of initial reactions, nor expose subjects to counterarguments, but even if they did, I am not sure how we would figure out that the subjects had engaged in reflection at all, let alone engaged in an admirable form of reflection. (Even if they changed their initial views—or if, in a between-subjects test, those put in the position designed to induce greater reflection had different views than those not exposed to purportedly reflection-inducing prompts—one would wonder whether the distinct views were a product of admirable deliberation rather than some

example, Cass R. Sunstein, *Hazardous Heuristics*, 70 U. CHI. L. REV. 751 (2003); and Cass R. Sunstein, *Moral Heuristics*, 28 BEHAV. & BRAIN SCI. 531 (2005). For a far fuller discussion of what is appealing and unappealing about Sunstein’s views, in light of more general debates about the perils and benefits of using heuristics, see Mark Kelman, *Moral Realism and the Heuristics Debate*, 5 J. LEGAL ANALYSIS (forthcoming 2013), available at <http://ssrn.com/abstract=2159521>.

44. See APPIAH, *supra* note 28, at 116-20, 125, 151-63.

45. Rawls discusses reflective equilibrium in JOHN RAWLS, A THEORY OF JUSTICE 48-51, 579 (1971). He states:

From the standpoint of moral philosophy, the best account of a person’s sense of justice is not the one which fits his judgments prior to his examining any conception of justice, but rather the one which matches his judgment in reflective equilibrium. . . . [T]his state is one reached after a person has weighed various proposed conceptions and he has either revised his judgments to accord with one of them or held fast to his initial convictions

Id. at 48. He criticizes direct reliance on case-by-case intuitions as well. *Id.* at 34-41.

unreflective reaction to an aspect of the prompt or experimental situation designed to elicit deeper consideration.)⁴⁶

3. *Unreflective intuitions: moral competency*

For a variety of reasons, I am skeptical that we can actually discover shared, universal moral reactions grounded in inborn competencies to represent moral problems identically across persons, in large part because I do not believe that our ultimate moral reactions are highly modularized. (That is to say, I don't believe that an output—a moral reaction—follows in a mandatory fashion from a small number of discrete environmental inputs in the same way that a reflex motion is modularized—a muscle movement follows inexorably from a small number of discrete inputs.) And I also believe that, even if I were wrong and we were to discover mandatory inborn moral reactions, roughly akin to reflexive movements or unlearned intuitive fears (e.g., snake phobia) or food tastes (e.g., taste for high-calorie sweets), they would still be entitled to little deference in particular cases, because they would at best reflect generally useful rules of thumb that might be, like any generalization, inapt (and we would want to explore more reflectively whether the reaction was appropriate in the situation)⁴⁷ and because they are likely to have evolved to solve particular social coordination problems that might or might not be the problems we most typically confront today. (The standard quickie story to get across this general

46. Kreps and I indeed found that people's responses to standard trolley problems do *differ* when put into situations that *could* be said to foster greater reflection (for example, by seeing cases side by side so that they must implicitly test whether their beliefs are consistent across cases, or by considering how a person they think of as a moral hero would respond to a dilemma). See Kelman & Kreps, *supra* note 29, at 37-40, 51-55, 62-65. But we do *not* make the claim that these arguably more reflective responses are normatively superior, or even that they should be viewed as more deliberative in any relevant sense. We make the far more modest claim that bottom-line reactions to the standard trolley problems are more *fragile* than those who believe such reactions are the product of strong immovable intuitions would imagine, and that certain substantive beliefs prove to be more fragile than others. Many people may *change* their view that it is permissible to divert the trolley onto a side track, killing one to save five on the main track, when they simultaneously consider that they are unwilling to push someone off of a drawbridge to block the trolley, and even more may do so when they are also exposed at the same time to cases that expose some of the standard objections to killing one to save many. But it is not clear at all that they have "reflected" more deeply on the divert-the-trolley case when they think about additional, arguably similar cases. They may well be better characterized as having been duped into making all of their responses consistent with an anti-pushing response that is *least* thoughtful and recalcitrant to analysis.

47. This is the view of those, like Sunstein, who believe that universal moral reactions do exist but that they are likely to be inapt to particular cases, just as judgment heuristics; the belief that events that are readily available to memory have been more common and are therefore, going forward, more likely to occur, may be wrong in particular cases. See Sunstein, *supra* note 41, at 1556-58. For a further exploration of the virtues and flaws of Sunstein's views, looked at through the lens of debates over the claim that heuristics frequently lead to error, see Kelman, *supra* note 43.

point about evolutionary lag is that our taste for sweets may have served us well in an environment in which calories were scarce but serves us poorly if it leads us nowadays to seek out foods that cause health-jeopardizing weight problems. Similarly, it might have made substantially more sense to show a strong preference towards members of one's in-group in a hunting and gathering society where non-kin were often competitors at best or overt threats at worst, and where needs for large scale social cooperation were minimal, than it makes to show such evaluative and normative preference for in-group members in a multicultural society with needs for mass cooperation. To put it bluntly, even if variants of xenophobia and racism made evolutionary sense, this should hardly count as an endorsement.)

Still, it is worth exploring briefly what an inborn moral competency might be and how we would identify one, and then, exploring further, why we might think such competencies (which are what I think we mean by an unreflective intuition) are especially important for normative law and philosophy scholars to identify. I think both what is clearest in this account and what is blurriest (both descriptively and in terms of normative implications) can be seen by looking at the fascinating work of John Mikhail. Mikhail has explored the relevant sort of universal, inborn competency largely in relationship to moral judgments about trolley problems.⁴⁸

48. Mikhail seeks to account for the purported universal competence to distinguish cases in which subjects judge the permissibility of diverting the trolley on to a side track, saving five on the main track but killing one on the side track, rather than push someone off the bridge to block the runaway trolley, and to distinguish cases in which a trolley is diverted on to a side track but will loop back to hit the five potential victims on the main track unless blocked by another man rather than killing a man before hitting a heavy object that will block the train from returning to the track. See, e.g., JOHN MIKHAIL, ELEMENTS OF MORAL COGNITION: RAWLS' LINGUISTIC ANALOGY AND THE COGNITIVE SCIENCE OF MORAL AND LEGAL JUDGMENT 77-122, 338-41 (2011). The competence that might be revealed by the purported tendency to distinguish these cases is first, the capacity to distinguish cases based on whether the batteries that occur are intended or unintended side effects of efforts to save the people on the main track and to measure how many distinct batteries are involved in the saving effort. If one reads Mikhail to claim *merely* that subjects have an inborn capacity to distinguish the cases—to see that the cases may be differentiated—his data supports (though hardly proves) that claim. If he means to prove either the claim that the bottom line reaction to the cases is universal, that any departure from the bottom line reaction that he expects (e.g., diverting is permissible, pushing impermissible) represents a performance error, or that the capacity to differentiate the cases determines reactions to the cases without further inputs as it would if moral reactions were modularized, then he has not succeeded. Bottom line reactions are not close to universal. They alter so much when people consider certain forms of counterclaims and troubling cases that they can be made to nearly converge on the same response (killing one is not permitted in either diverting or pushing cases); and they are radically more complex than first-order judgments of permissibility and impermissibility suggest (roughly half think the requirement to divert is mandatory, another quarter or so think it merely permissible; the judgment that diverting is permissible is both more weakly held—that is, seen to be a closer question—than the judgment that pushing is forbidden, and it is more likely to be abandoned when the person who would be killed by diverting is related to the person considering pushing or is more thickly identified). For a far fuller discussion of

The first question one must ask is how we might identify the sorts of intuitive moral competencies we might care about and distinguish them from mere *opinions* on issues (that can be learned—and presumably learned differently—and adjusted through reflection, and that are likely to differ across persons). Mikhail's descriptive claims draw heavily on his decision to employ a linguistic analogy in thinking about why we might be said to have a natural sense of justice: all people, everywhere, are born with a "moral competence" that closely parallels the linguistic competence that permits people to learn *a* language.⁴⁹ While the details of moral rules may differ in some respects, the existence of intuitive competencies that *matter* to us for these purposes does guarantee that the capacity to represent moral problems as possessing certain features is shared⁵⁰ and that a significant number of moral reactions, grounded in these representational features, are likely to be universal rather than culturally specific (just as certain grammatical building blocks are purportedly universal across the range of languages).⁵¹ What is most critical in identifying whether we are observing this form of intuition (rather than, say, a learned reaction) are that these sorts of intuitions have the following traits that map closely to linguistic competence: people making judgments of permissibility and obligation grounded in these sorts of intuitive basic competencies are able to do so without being able to explain *why* they have made them;⁵² people making these representations and corresponding judgments could not possibly have learned to make them by generalizing from more particular moral rules they have been exposed to (this is the familiar "poverty of the stimulus" argument);⁵³ and the moral *system* grounded in these competencies is built from a small set of building blocks (e.g., all people naturally distinguish conventional from moral rules and believe

these empirical hesitations about one view of what Mikhail is claiming, see Kelman & Kreps, *supra* note 29, at 60-61. But, for now, I want simply to *assume* that it were true that there was a universal, bottom line, invariant, determining competence to distinguish diverting and pushing cases and speculate about why that might be *more* relevant to normative lawyers and philosophers than a more considered or reflective judgment.

49. The particular languages people learn are obviously not identical; in that sense, languages are conventional and diverse rather than universal. But what some—notably Noam Chomsky—have argued was that the capacity to learn *a* language is grounded in an innate and universal human competence to recognize a host of things necessary to speak any conventional language, for example, the distinctions between sentences and pseudo-sentences, ambiguous and clear expressions. In much the same way, the fact that we have the competence to acquire morality does not dictate the content of each and every moral rule, nor does it guarantee that we will "perform" moral judgment-making properly in all settings (e.g., regardless of how the problem is framed or regardless of the inexorable tendency to make errors). See MIKHAIL, *supra* note 48, at 51-55.

50. *Id.* at 114-17.

51. *Id.* at 299-303.

52. *Id.* at 19-21, 83-84, 323, 330.

53. *Id.* at 17, 69-71.

that if an act or omission is not forbidden, then it is permitted, rather than that all permitted acts must be specified).⁵⁴

It is imperative to note, before discussing why these sorts of competency intuitions may have particular normative force, that the responses to most vignettes in the empirical psychology literature, including those proffered by the *Killing in War* researchers, do *not* appear to elicit these sorts of reactions. Nothing in the experiments that the *Killing in War* authors report suggests that subjects could not defend their judgments about the permissibility of particular attacks, that they would have had opinions about the issues the experimenters raise prior to being taught lessons transparently relevant to the prompts, or that they can readily categorize rules about killing soldiers only, for instance, as moral rather than conventional rules that are a product of convention or treaty, born of either the imposition of the desires of internationally dominant nations, or some consequentialist judgment shared by representatives of potentially warring nations.⁵⁵ So *if* the only sort of intuitions we should care much about in drawing normative conclusions from popular intuitions are ones with this profile, my strong suspicion is that the intuitions in *Killing in War* are irrelevant. But it is important to draw a prescriptive lesson here too: *if* these sorts of intuitions are the only sort that provide true normative support for a proposition and *if* a particular author wants to use vignette-response experiments to bolster a normative claim, it is imperative that the experimenter demonstrates that she is eliciting only this sort of intuition.

As I said, I am extremely skeptical that even if we could identify inborn competencies to represent certain problems in particular ways that those representations would dictate how we felt a moral quandary ought to be resolved, and equally skeptical, ultimately, that the beliefs we form that are somehow formed more effortlessly or inexorably are morally superior. But it strikes me as worth giving a brief outline of the case that such beliefs are entitled to special moral deference, because, in the final analysis, I am even more skeptical that there is a strong case for believing that any *other* intuitions have intrinsic normative force.⁵⁶

Mikhail rejects, quite rightly in my view, one sort of argument that might be offered in defense of the proposition that observed universal moral rules

54. John Mikhail, *Universal Moral Grammar: Theory, Evidence and the Future*, 11 TRENDS COGNITIVE SCI. 143, 143-45 (2007).

55. Once more, one reason that researchers may typically make few claims that distributive justice intuitions are of much normative weight is that they typically have none of these traits. See *supra* note 15.

56. Recall that it is a different (albeit perfectly important) issue whether, for instance, we should adopt those moral rules that match intuitions because such rules will be more readily accepted and obeyed, but it should be noted, in this regard, that if it is indeed hard to shed certain reactions to moral problems, certain rules may be hard to imbed. Resistance at the cognitive level to understanding a counterintuitive proposition might well translate into resistance at the behavioral level to implementing counterintuitive rules.

must be correct: the idea that those committed to this sort of “moral realism” must believe that people are perceiving the same moral rules because they are perceiving or *discovering* rules that preexist in the external world. Such “external rules” might be established by external, presumably divine authority,⁵⁷ or, more interestingly, be thought of as external because the rules are, akin to physical laws, the only set of rules that could conceivably govern functioning social relations. Mikhail clearly states that he is not committed to any sort of “mind-independent moral reality,” and his views seem to me representative of those drawn to this sort of intuitionism. In his view, the source of whatever universalized moral injunctions we observe is within our minds—the cognitive mechanisms that permit us to acquire moral beliefs.⁵⁸

In terms of the normative significance of Mikhail’s claims, it would obviously strengthen the idea that our moral “intuitions” have at least presumptive normative force if, first, the intuitions that we had developed were maximally adaptive because of their impact on social/legal organization,⁵⁹ and second, if

57. Obviously, if one believes that there are divinely implanted moral sentiments, discovering what they are would plainly be of normative moment.

58. See, e.g., MIKHAIL, *supra* note 48, at 220-21. This claim is interestingly related to claims about the significance of (whatever) universalism of beliefs we might discover. One can imagine a theorist who thoroughly rejects Mikhail’s claims about cognitive moral competence, believing instead that moral beliefs are some mixture of learned conventions and the product of independent reasoning performed by relatively fully cognitively flexible minds. Such theorists could still believe that we would observe at least some “universals” if they believed that only certain moral “systems” would survive and permit adequate levels of social cooperation to sustain a culture: universalism would result, in this view, from some combination of “cultural evolutionary processes” (survival of only those *groups* following certain norms) and/or the convergence of rational cognitively generalist minds on the small set of rules that were actually workable.

At the same time, one might imagine that certain sorts of diversity in these unreflective moral attitudes would be expected to the extent that it was adaptive for one subgroup, but not another, to follow a particular code. If one accepts (solely for sake of argument and illustration, in my view) the mainstream pop evolutionary psychological account of the distinction between male and female sexual morality and behavior—men find certain sorts of promiscuity and even coercive sexuality acceptable because gene replication ends are advanced by maximizing the number of females impregnated, while women develop a sexual morality grounded in fidelity obligations and norms permitting them to manifest pickiness about partners—one would *expect* gender diversity in intuitions on moral topics involving sexuality and one would expect that each moral code is adaptive, though they cannot be fully reconciled nor is one more unambiguously preferable from the vantage point of maximizing survival *rates*, even if one accepted that survival rate maximization were a good normative measure.

59. I am dubious whether this sort of claim could be sustained if thinking about some of the rules that Mikhail thinks are imbedded—for example, rules about double effect. But, of course, this skepticism would have to be extended to the implicit evolutionary claims in Mikhail’s work that a certain Universal Moral Grammar dictating or facilitating a particular set of moral rules produced direct reproductive advantages by insuring certain sorts of social organizations with certain sorts of conduct rules stabilized. The claim seems even more attenuated if the (putative) moral competency were merely an explicable side effect of a cognitive structure that produced entirely distinct inclusive fitness gains. For instance, if one im-

being maximally adaptive were either equivalent to being normatively preferable or at least a strong factor to be weighed in judging the acceptability of a norm.⁶⁰ Neither proposition is transparently correct, either in my mind or, as far as I can discern, in Mikhail's, but surely, each idea is worthy of at least some serious consideration. One might hesitate to extol the adaptive, even as a first approximation, merely because it is a commonplace in evolutionary biology that traits that evolve need not be the best traits imaginable to solve a problem. If nothing else, there are developmental biological limits on the range of plausible mutations. And it is not clear why we should judge the ultimate moral acceptability of a belief solely in terms of the fit between that belief and an individual's chances of passing on his genes. Inclusive fitness and moral acceptability may simply be unrelated to one another. It is not precisely clear what if anything Mikhail thinks follows normatively were we to accept the idea that a fuller Universal Moral Grammar (UMG) might someday be specified and understood.

It is plausible that Mikhail *does* believe—and that it is a reasonable belief—that our best understanding of what an ideal moral rule might be is the rule that emerged that permitted the level of social cooperation that sustained human life over time. There is a strong adaptationist version of this, and a weaker Burkean version: how can we be vain enough to believe that any of us is smart enough as an individual to imagine a set of working moral rules that does better than the ones that emerged through some (metaphorical equivalent of) competition among rules? But it is also possible that what Mikhail might ultimately argue, instead, is that the question of whether a moral principle or practice generated by the UMG is by virtue of that a “good” or “acceptable” argument is in some ways senseless. This might be true in much the same way as it would seem rather senseless to ask whether the linguistic structures we can generate given the linguistic competence represented by the universal grammar are “good” linguistic structures. If we believe that the moral domain is *defined* as the “output” of the morality-creating module, the products of the UMG *are* human morality.

There is yet another argument that I tentatively suspect comes even closer to Mikhail's view, and may seem more compelling still. If it were the case that only certain moral rules could be readily learned and readily attract agreement, it would be of some weight in evaluating those rules. The degree of weight is

agined that representing trolley dilemmas in ways consistent with the Doctrine of Double Effect were a mere by-product of some cognitive capacity to make causal inferences about mechanistic relationships important to our capacity to solve completely distinct problems (in hunting, in avoiding predators, whatever) with no moral content at all, it could still be said to have adaptive advantages, but not because of its role in cementing a well-functioning social world. I am always dubious about our capacity to isolate the single critical feature of a cognitive representation.

60. For an argument that commonplace moral reactions are likely adaptive and normatively compelling for that reason, see Paul H. Robinson et al., *Origins of Shared Intuitions of Justice*, 60 VAND. L. REV. 1633 (2007).

quite ambiguous: it is possible that the fact that a rule is acquired effortlessly and attracts consensus is, in Mikhail's view, merely a mild factor in its favor or a near-killer argument for its acceptability. As alternative rules become harder and harder to adhere to, the argument that "ought follows is" gets stronger: one gets closer to arguing that the alternative rules can be taken off the table simply because getting people to treat them as the sorts of moral rules that they can comprehend (let alone feel have a force beyond the force inherent in any command backed by threats) is as infeasible as making a rule that they flap their arms and fly to work in order to avoid emitting greenhouse gases.

CONCLUSION

There are plenty of good reasons for a policymaker to care whether legal rules seem immediately morally acceptable to those who will be subject to the rules: people may learn rules that they intuit should govern their conduct more readily than they would learn counterintuitive rules; they might behave in accord with the rules they believe should exist no matter what rules *do* exist; they might well obey such intuitive rules more readily even when they do not fear that noncompliance will lead to penalties; they might find a legal system more legitimate if positive law typically matched beliefs. If legislation should respond to popular will, vignette response gives us one sort of information about one relevant sort of public opinion.⁶¹ There are plenty of reasons too for scholars interested in cognition to study responses to vignettes in order to learn more about moral reasoning and morality acquisition, just as there are reasons to study the circumstances in which people make certain suppositions about whether they know a true fact, whether or not popular beliefs about what we truly know are of any use to epistemologists figuring out when beliefs are actually warranted or not.

It is not nearly so obvious that we should care much about how subjects respond to vignettes if our goal is to figure out what is normatively best. There is a standard experimentalists' story: law should (at least sometimes) require what is morally desirable; what is morally desirable can best be ascertained by figuring out moral principles that comport with or generate the moral intuitions people actually have; "intuitions about intuitions"—armchair philosophers' guesses about what commonplace intuitions actually are—may well be wrong. But the story is problematic in many ways. It is not at all clear that normative moral philosophers actually do, as a descriptive matter, care much about intuitions, but what is more important for these purposes is that it is not at all clear that

61. The arguments I explored about whether reflective intuitions are more worthy of consideration than unreflective competencies might be resolved differently if one were interested, above all, in democratic responsiveness. The debates about whether we should attend more to the opinions generated by a deliberative democratic process rather than unreflective opinions is significantly different than the questions we tried to answer here about whether arguably adaptive competencies were entitled to particular deference.

they *should*. Typically, vignette responses are unduly diverse and can scarcely be distinguished from learned beliefs, or the product of the very sort of normative arguments that philosophers could readily make (and arguably make better) without caring which arguments have proven more persuasive to more listeners. There is an interesting case to be made that certain intuitions—unreflective, unlearned, and universal ones whose persuasive force confounds the respondent—merit more normative attention. But it is unclear that the capacities that have these qualities (above all, unlearned and opaque) actually determine bottom-line moral attitudes, and it is also unclear whether they would be worthy of moral deference even if they did, both because they may simply be reactions that work in the bulk of cases we confront, but misfire in atypical ones that may be the source of legal controversy, or because they may be reactions that were appropriate to the problems that most typically confronted human beings over the many millennia in which we evolved, but map poorly to the moral problems of modernity. If, though, one is drawn to the idea that moral judgments grounded in unreflective, unlearned competencies have special normative force, it is important that experimenters do more to help us identify when they are reporting a reaction likely to be grounded in such competencies and when they are merely reporting commonplace opinions.

