Groundwork for a New Moral Epistemology

Marcus Arvan
(University of Tampa)

> For a while I thought: 'Well, the arguments are right, capitalism is the best system, but only bad people would think so. Then, at some point, my mind and my heart were in unison.'
>
> -- Robert Nozick, to a journalist[1]

> [I]t is not profitable for us at present to do moral philosophy…until we have an adequate philosophy of psychology, in which we are conspicuously lacking.
>
> -- EM Anscombe[2]

Philosophy in the Western tradition has always been concerned with arguments. Yet arguments involve premises, and it is often a matter of deep and lasting disagreement which premises are true. Consider for example the longstanding debate between libertarians and liberal-egalitarians in political philosophy. In *Anarchy, State, and Utopia*, Robert Nozick defends a libertarian political theory based on the premise that people have more or less absolute moral rights to life, liberty, and property – rights that can never be permissibly violated for the good of others. Libertarianism's opponents typically reject these premises. Liberal-egalitarians, for example, tend to argue that political principles must be founded either on some kind of overlapping social consensus[3], ideal of reciprocity[4], or the mitigation of brute bad luck.[5] Yet

* I thank David Morrow, Trent Dougherty, two anonymous referees, and audiences at the 2012 Rocky Mountain Ethics Congress and 2011 Conference on the Normative Implications at the University of Alabama at Birmingham for their helpful and challenging comments. I also thank Florian Cova for his work editing this special issue.

[1] As reported in: http://www.slate.com/articles/arts/the_dilettante/2011/06/the_liberty_scam.html (accessed 31 March 2013).

[2] G.E.M. Anscombe, « Modern Moral Philosophy », in *Philosophy*, 33/124, 1958, p. 1.

[3] See e.g. J. Rawls, *Political Liberalism*, Columbia, Columbia University Press, 1993.

[4] J. Rawls, *A Theory of Justice*, Cambridge MA, The Belknap Press of Harvard University Press, 1993.

[5] See e.g. R. Dworkin, *Sovereign Virtue*, Cambridge MA, Harvard University Press, 2000.

libertarians typically reject *these* premises. Libertarians and liberal-egalitarians thus seem to fundamentally disagree over premises. Next, consider Utilitarianism, Kantianism, and Virtue Theory in moral philosophy. Utilitarians tend to defend their theory on the premises that (a) happiness is each person's good, and (b) moral rightness is a matter of impartially pursuing the good.[6] Kantians (including, of course, Kant himself), on the other hand, tend to argue from the premise that morality is a matter of conformity to categorical norms of practical reason[7] or the pursuit of unconditioned goodness.[8] Finally, virtue theorists typically argue from premises that (a) each person pursues their own happiness or flourishing, and (b) particular character traits (i.e. the virtues) are conducive to happiness or flourishing.[9] Here again, the various sides seem to simply disagree over premises. These kinds of disputes raise the inevitable question: are there really no deeper grounds – premises that all sides to these debates can accept the truth of – for resolving these disagreements?

This paper aims to show that such grounds may indeed exist. It argues that recent empirical research tentatively supports a new kind of *empirically-informed moral-virtue epistemology* – an epistemology of evaluating disputed premises in moral and political philosophy by reference to empirically-observed relationships between (a) judgments about moral premises, and (b) personality, character, and behavioral traits, the moral value (or valence) of which *all parties* to such debates agree upon. More exactly, this paper defends the following two claims:

---

[6] See e.g. J. Bentham, *An Introduction to the Principles and Morals of Legislation*, Amherst NY, Prometheus Books, 1789, chap.1; J.S. Mill, *Utilitarianism*, reprinted in G. Sher, *John Stuart Mill: Utilitarianism and the 1868 Speech on Capital Punishment, 2nd edition*, Indianapolis/Cambridge, Hackett Publishing, 2002, chap. 4; and more recently R.M. Hare, *Moral Thinking*, Oxford, Oxford University Press, 1981; P. Singer, *Practical Ethics*, Cambridge, Cambridge University Press, 1979; and R. Brandt, Richard, *A Theory of the Good and the Right*, Oxford, Clarendon Press, 1979.

[7] I. Kant, *Groundwork of the Metaphysics of Morals*, Cambridge, Cambridge University Press, 1997.

[8] See e.g. B. Herman, *The Practice of Moral Judgment*, Cambridge, Harvard University Press, 1996.

[9] See e.g. J. Annas, « Comments on John Doris' Lack of Character », in *Philosophy and Phenomenological Research*, LXXI/3, 2005, pp.636-642; and R. Hursthouse, *On Virtue Ethics*, Oxford, Oxford University Press, 1999, chap. 8.

- *Epistemological Claim:* Prevailing epistemological norms in moral and political philosophy entail that we ought to aim, as far as possible, to epistemically privilege – in argument and theory-construction – moral premises endorsed by those among us who have the "best moral compass", as defined by reference to personality, character, and behavioral traits commonly recognized, by all parties to the relevant moral-political debates, to be morally good, bad, right, and wrong.[10]
- *Empirical Bet:* There are good reasons to believe that empirical philosophical-psychological research may be able to provide real (albeit imperfect) inductive evidence regarding (a) which individuals among us have the "best moral compass", and (b) which moral premises those people endorse.

In short, this paper argues that insofar as a strong case can be made for both of these claims, philosophers and psychologists have compelling epistemological reasons to engage in a vast new research program: an *experimental ethics* utilizing empirical research on personality, character, and overt behavior to inform our judgments about the types of premises it is legitimate to invoke in moral and political philosophy; a research program which *may* be able to provide widely acceptable grounds – grounds that *all* parties to disputes can accept – for resolving fundamental disagreements about the truth of disputed premises in moral and political philosophy.

§1 of this paper defends Epistemological Claim. §2 then defends Empirical Claim. Finally, §3 raises and responds to a number of potential objections.

[10] One might ask: why think it is necessary for all parties to debates to agree on which traits are good/bad? To which I answer: a primary aim of the new moral epistemology I propose is to *forge greater agreement* in moral and political philosophy, enabling people in these fields to move beyond entrenched disputes, in much the same way that the development of the scientific method enabled investigators of the physical world to move beyond ancient speculative claims (e.g. everything is made of water, as Thales claimed; everything is air, as Anaximenes claimed) to broad consensus on the nature of physical reality (e.g. the Standard Model of physics). I thank an anonymous referee for raising this question.

I. The Case for "Epistemological Claim"

When two or more individuals disagree over moral premises – premises that figure into arguments in moral and political philosophy – can there be any good epistemic grounds for thinking that the premise(s) one person finds attractive are more likely to be true than the premise(s) another person finds attractive? The most obvious way to try to proceed is to find some still more fundamental premise(s) that the parties to the dispute can agree upon to resolve their disagreement. So let us try.

Return to the two types of disagreement over premises mentioned in the introduction to this paper: disagreements between libertarians and liberal-egalitarians in political philosophy, and disagreements between Utilitarianism, Kantianism, and Virtue Theory in moral philosophy. As we have already seen, the parties to these disagreements appear to disagree over premises. Utilitarians begin from the premise that moral rightness is a matter of impartially promoting the good; Kantians begin from the premise that morality is categorical in nature; and virtue theorists begin with premises regarding human happiness or flourishing. Now, although these different camps disagree over these premises, are there really no further premises, quite aside from the ones they disagree about, that all sides can accept as true? Surely not: libertarians, liberal-egalitarians, utilitarians, Kantians, and virtue-theorists – all of the disputing parties cited in this paper thus far – all agree that things like lying, cheating, and stealing are (generally) morally wrong and bad, and things like honesty, helpfulness, conscientiousness, etc., (generally) morally good. Although there are of course some people on the philosophical fringes (e.g. moral skeptics, etc.) who are willing to deny the truth of even these moral judgments, let us set these perspectives aside for now (we will return to them later). The important thing for now is that there is overwhelming agreement among mainstream moral and political philosophers on these issues.

Now consider the epistemic situation of a person making a judgment

about the truth of some disputed moral premise. Consider, for example, Robert Nozick's judgment in Anarchy, State and Utopia that people have natural moral rights to life, liberty, and property which can never be permissibly violated for the good of others. Here is something that I propose all of us – including Nozick – can agree upon: if Nozick is to be an epistemically responsible agent, he should be prepared to (A) ask whether there is anything about him, psychologically, that might have morally corrupted his judgment that P is true, and, if any such corrupting propensity is indeed discovered, (B) discount (or lessen) his level confidence in the truth of that premise. The reason I think we can all agree upon this is simple: individuals' judgments about moral premises clearly can be morally corrupted by their psychology, in ways that ought to lead them to epistemically discount their confidence in those judgments. An obvious case here is a Nazi. A Nazi might firmly accept the premise, "the Aryan race morally ought to dominate the world." However, it is plain to the rest of us that the premise is false, and that the Nazi only believes it is true because they have a distorted moral perspective. The Nazi ought to doubt (and indeed, reject!) their premise because they believe it on morally vicious grounds: their hatred of other races. An even more obvious case is a psychopath. Psychopaths appear to sincerely judge that there is nothing wrong at all with lying, cheating, stealing, or killing. Why? Answer: because they are psychopaths.

My next claim is this: philosophers and ordinary people commonly *recognize* that a person's "moral compass" (i.e. their personality, character, and behavioral traits) can epistemically improve or distort judgments about the truth of moral and political premises. The brutal dictator may "see nothing wrong" with killing millions of people because they are a morally corrupt person. The psychopath may fundamentally see nothing wrong with torturing people because they are psychopathic. The callous person may see nothing wrong with helping an old lady cross the street because they are callous. Conversely, a kind person will judge it "obviously right" to help an old lady across the street because of her kindness. Etc.

Now, insofar as we commonly recognize this – that an individual's moral compass (i.e. their psychological make-up) can improve or distort their

judgments about moral and political premises – we seem to share the following epistemological premise in common: we ought to (A) investigate the ways in which different psychological characteristics improve or distort judgments about moral and political premises, and (B) assign greater epistemic credence to moral premises found attractive by those with the "best moral compass" (i.e. the person with the morally best psychological make-up) than to those premises found attractive by those with more morally corrupt psychological profiles. Let us investigate these ideas in more detail.

Ad hominem arguments are fallacious in many circumstances. In most domains, a morally bad person's beliefs can be as, or even more, truth-apt than a good person's beliefs. For example, suppose Jones is morally good person, but Jones knows little about particle physics. Jones' friend, Physicist, may be a morally bad person and yet know far more about particle physics than Jones. It is plainly fallacious to suggest that Physicist's judgments about premises in particle physics should be distrusted because he is a morally worse person than Jones. Still, ad hominem arguments are not obviously fallacious in the case moral judgments (including those that figure into political philosophy). When a psychopath claims, "There is nothing wrong with wantonly murdering people", we are apt to respond: "Only a sick person could believe that." We reject the psychopath's moral premise as false because we recognize that they a moral monster. We recognize that their badness of character makes them sensitive to, and find attractive, moral premises that only a bad person could find attractive. Thus, when they make moral assertions that strike us as monstrous, we simply reject their claims on those grounds alone: as monstrous.

Notice, furthermore, that there is a mainstream approach to moral philosophy that appears capable of legitimizing the epistemic value of *ad hominem* arguments of this sort: virtue theory. Virtue theorists wish to understand some, or all, of morality in terms of the virtues of character. Rosalind Hursthouse, for example, defends an analysis of morally right action that reduces right action to the characteristic actions of the virtuous agent. She argues that the morally right thing to do in a given circumstance is that which

the virtuous person characteristically would do in that circumstance.[11] Notice what follows straightforwardly from this analysis of right action: the *morally right* premises to use in moral and political argument and theory are the premises that the *virtuous* person judges to be true. Conversely, we ought to reject the bad person's judgments about disputed moral premises as likely corrupted by *psychological moral deficiency*.

Now although I have just used a virtue-theoretic definition of right action to justify this epistemological principle – the principle of prioritizing the good person's judgments about moral premises over the bad person's – it is already (if only implicitly) *standard practice* in moral and political philosophy. For notice: when moral and political philosophers construct arguments and theories, they do not begin with the psychopath's, Nazi's, or immoralist's (the person who endorses immoral behavior) favored premises. No, it is standard practice to begin with premises that "we" – ordinary, decent-minded, philosophically reflective people – believe to be true. So, for example, when Robert Nozick bases his libertarian political theory on the premise that human beings have natural moral rights to life, liberty, and property, he is surely not assuming that the psychopath, Nazi, or immoralist would accept the premise (they presumably wouldn't); he is assuming that *we – morally better people* – will accept it, and that it is epistemically appropriate to invoke the premise *for that very reason*: the fact that, *qua* morally better people, we are *better judges of moral premises* than those other people. Similarly, liberal-egalitarians in political philosophy premise their theories on appeals to "our" judgments about fairness, reciprocity, overlapping consensus, etc. – premises they assume *people like us* will accept as true (not the Nazi, psychopath, or immoralist). Here again the assumption is that we are better judges of the truth of moral premises than the psychopath, Nazi, or immoralist *because we are better people* The same assumption is also standard practice in moral theory: Utilitarians appeal to "our" assumptions about happiness and impartiality, Kantians appeal to "our" judgments about practical reason, categorical normativity; etc. Notice, again, that *no* attempt is (ordinarily) made to justify these starting-points to Nazis, psychopaths, or immoralists.

[11] R. Hursthouse, *On Virtue Ethics*, Oxford, Oxford University Press, 1999, p. 28.

It is crucial to emphasize here just how routine and ubiquitous this epistemological move is in moral and political philosophy. A vast array of influential arguments in moral and political philosophy are predicated on the assumption that some of us – ordinary, everyday "decent" people – are *epistemically better judges of moral premises* than other people precisely on account of our better moral character. Consider, for example, Judith Jarvis Thompson's famous "Violinist Case" argument concerning the ethics of abortion. Thompson's argument is based on an appeal to how "we" judge particular case: the case of a person kidnapped in the night and hooked up by a series of series of medical tubes to a famed dying violinist to save the violinist from a fatal kidney ailment. Thompson argues that if you were the person hooked up to the violinist and someone told you that you have no moral right to unhook yourself, "you would regard this as outrageous."[12] Thompson then argues from this alone – from "our" moral judgment in the case – that whatever else the moral right to life is, it is not a moral right to depend on another person's body for life. Of course, not everyone shares this judgment. A religious fanatic might judge the case differently (viz. life is more sacred than a person's right to their own body), and presumably a psychopath would *definitely* disagree (they might even say, "Not only does the violinist not have a right to your body. If you want to slit his throat, go right ahead!"). What Thompson is tacitly assuming, then – in making her argument – is that *we, her readers, the ones who judge the case the way she does, have morally better dispositions, a better "moral compass"*, than those who judge the case differently.

This is far from an isolated case. Arguments in applied ethics and political philosophy on issues as diverse as euthanasia[13], affirmative action[14], duties of the affluent to give to charity[15], homosexuality[16], war[17], the basic

[12] J.J. Thompson, « A Defense of Abortion », in *Philosophy and Public Affairs*, 1, 1971, p. 818.
[13] See e.g. J. Rachels, « Active and passive euthanasia », in *Bioethics: An Introduction to the History, Methods, and Practice*, London, ones and Bartlett Publishers, 1997, pp. 1-82.
[14] See e.g. L. Pojman, « The Case Against Affirmative Action », in *Ethics in Practice, 3rd edition*, Malden, Blackwell Publishing, 2007, pp. 473-484.
[15] See e.g. P. Singer, "Famine, affluence, and morality", in *Philosophy & Public Affairs*, 1, 1972, pp. 229-243.
[16] See e.g. J. Corvino, « Homosexuality and the Moral Relevance of Experience », in *Ethics in Practice, 3rd edition*, Malden, Blackwell Publishing, 2007, pp. 298-308.

moral and political rights of human beings[18], etc. all typically begin with appeals to how "we" judge particular cases and principles. And again, consider the longstanding debate between different moral and political theories. Such theories are commonly defended by their proponents in large part on the basis of how well they embody or cohere with "our" considered moral judgments.[19] Utilitarianism is thought to embody "our" judgment that happiness matters, and that things like lying, cheating, and stealing are wrong because they tend to cause *unhappiness*. Similarly, Kantianism is thought to embody "our" judgment that people are not to be used like objects, but rather *respected*. Etc.

Now, it is not as though philosophers *never* confront the Nazi or the immoralist (engaging in moral argument with a true psychopath is impossible, for obvious enough reasons). There have many attempts to refute immoralism (see e.g. Plato's *Republic* and Hobbes' *Leviathan*, among others), as well as of course moral arguments against Nazism. The problem, however, is that it is hard, if not impossible, to refute the immoralist, psychopath, or Nazi on the basis of moral premises *they* accept (since, again, they tend to accept such different premises than we do).[20] Consequently, it has simply become common practice in moral and political philosophy to simply begin with premises that "we" find attractive, and set aside the Nazi, immoralist, and psychopath as unreasonable and outside mainstream moral conversation. The background assumption is that if we cannot definitively refute the immoralist, etc., on premises they are apt to accept, *the epistemically appropriate thing to do is to simply begin with the premises that "we" – morally better people – believe to be true*.

Now, it may be objected that this is terrible epistemological development in the profession – that we *shouldn't* just assume morally better individuals' judgments about moral premises (i.e. our judgments) are any more

[17] See e.g. D.P. Lackey, « Nipping Evil in the Bud: The Questionable Ethics of Preventive Force », in *Ethics in Practice, 3rd edition*, Malden, Blackwell Publishing, 2007, pp. 715-725.
[18] Again, see R. Nozick, *Anarchy, State, and Utopia*, USA: Basic Books, 1974. Also see H. Shue, *Basic Rights*, Princeton, Princeton University Press, 1980.
[19] See e.g. M. Timmons, *Moral Theory: An Introduction, 2nd ed.*, Lanham, Rowman and Littlefield Publishers, 2013, pp. 169 and 239.
[20] See e.g. R. Joyce, *The Myth of Morality*, Cambridg, Cambridge University Press, 2001, chap. 2 for a nice discussion.

likely to be true than the judgments of psychopaths, Nazis, or immoralists. But again, the problem is that there often seems no better way to go, at least if moral and political philosophy are to be at all productive. If there is simply nothing we can do to convince the psychopath, Nazi, or immoralist that premises are false – and they often simply do begin with different premises – then, if we are to engage in productive (as opposed to skeptical[21]) moral and political philosophy, it seems entirely appropriate to assume that "our" moral judgments – *qua* judgments of morally better people – are more likely to be true.

I should probably emphasize that I am not claiming that there are *no* other philosophical options than this standard practice.[22] One can argue, as R.M. Hare did for example, that moral terms have logical and semantic properties that normatively require universalizability and prescriptivity – in which case one might maintain that the immoralist, psychopath, or Nazi misunderstands moral language itself.[23] Or, one might maintain that moral terms are emotive in meaning[24] – and, perhaps, as expressive of the moral sentiments of an impartial spectator[25] – in which case one need not convince the immoralist, psychopath, or Nazi at all in order to vindicate moral norms. Etc. My claim in this paper isn't that there are no other philosophical options aside from casting aside the premises of the immoralist, Nazi, and psychopath on account of our moral goodness and their moral badness. My claim is *merely* that it is standard epistemic practice to prioritize our premises over theirs *on*

[21] We cannot, for obvious reasons, engage the moral skeptic here – though there are very notable defenses of the position (see e.g. R. Joyce, *The Myth of Morality*, Cambridge, Cambridge University Press, 2001). If the moral skeptic is correct, moral and political philosophy as a whole are *epistemically* misguided enterprises (though, perhaps, instrumentally useful). Accordingly, let us simply assume throughout the present inquiry – along with many in our discipline – that the moral skeptic is incorrect. Because a full and proper evaluation of moral skepticism must occur elsewhere, we shall simply set skepticism aside, and aim to describe and assume the standard epistemological norms endorsed by non-skeptical moral and political philosophers.

[22] I thank an anonymous referee for pressing this worry.

[23] See R.M. Hare, *The Language of Morals*, Oxford, Clarendon Press, 1952.

[24] See e.g. C.L. Stevenson, « The Emotive Meaning of Ethical Terms » in *Mind*, 46, 1937, pp.14-31.

[25] See e.g. D. Hume, *An Enquiry Concerning the Principles of Morals*, Oxford, Oxford University Press, 2007.

account of our moral goodness and their moral badness. And I think it is clear, from what has been discussed so far, that this is indeed standard practice.

If this is right – if it is standard practice in moral and political philosophy to epistemically privilege moral premises believed by better people (viz. "our considered judments") over moral premises accepted by morally worse people (e.g. the immoralist, psychopath, Nazi, etc.) – what follows? Let us begin with a toy case.[26] Recall Nozick's premise that human beings have natural moral rights to life, liberty, and property that can never be permissibly violated for the good of others. Call this premise *P*. Nozick judges *P* to be true, or at least likely to be true, and therefore legitimate to invoke in philosophical argument and theory construction. Now consider another individual – let's call him "Rawls" – who believes a very different premise to be true: one inconsistent with Nozck's premise P. Suppose Rawls' premise is: "Political principles morally ought to be founded in an overlapping social consensus." Call this premise *Q*. Since an overlapping social consensus could settle on something other than Nozick's idea that people have natural moral rights, *Q* entails *not-P*. Rawls and Nozick fundamentally disagree over moral premises.

Let us now return to the norm of assigning greater epistemic credence to the moral judgments of morally better individuals (e.g. decent, philosophically reflective people like us) over the moral judgments of morally worse individuals (e.g. the Nazi, the immoralist, the psychopath) – an epistemic norm which I have argued is not only (A) embodied in standard practices in the profession, but also (B) entailed by a virtue-theoretic analysis of right action. Suppose, next, that Rawls, Nozick, and almost the rest of us all agree on the moral value of particular personality, character, behavioral traits, and actions: for instance, that lying, cheating, and stealing are wrong actions; that liars, cheats, and murders are bad people; that being deceptive, being disposed to cheat, being disposed to steal, being disposed to callously disregard the feelings of others, being disposed to cruelty, etc. are moral vices; and conversely, that conscientiousness, kindness, helpfulness, fair-mindedness, etc. are moral virtues. Next, suppose it were an *empirical fact* that Nozick both had a lot of the

[26] Readers should be forewarned that the case is intended to be a "toy case" – one that is not intended to be empirically plausible. The aim is to tell a metaphysically possible but empirically implausible story to illustrate a *general point about epistemic norms*.

bad psychological traits and actually tends to act badly (he tends to lie, cheat, etc.), whereas Rawls had a lot of the good traits and tended to act in good ways (he tends to be honest, help others, etc.). Finally, suppose it were an empirical fact that Nozick's morally bad traits and dispositions *cause* him to make judge that *P*, whereas Rawls' morally good traits and dispositions were causally responsible for his judgment *Q*. If the world really were this way *and we obeyed the epistemic norm I have argued is standard practice in moral and political philosophy*, it would follow that we ought to epistemically privilege Rawls' premise *Q* over Nozick's premise *P* in argument and theory construction. Moreover, since Rawls and Nozick *both agree themselves* that lying, cheating, stealing, etc., are morally bad and wrong, and honesty, kindness, etc. are morally good, etc., if *they* were presented with empirical evidence of the above facts (i.e. the facts linking Rawls' judgments to good traits and behavior and Nozick's judgments to bad traits and behavior), then, epistemically speaking, *they* ought to assign greater epistemic credence to Rawls' *Q* than Nozick's *P*. Nozick, in particular, *if he were to obey standard epistemic norms*, should be led by the evidence to *have far less confidence* in the truth of his favored premise *P* and more confidence in Rawls' premise *Q*. After all, Nozick would have *empirical evidence* that his belief in *P* – as "obviously true" as *P* might seem to him – is the result of kind of *moral corruption of his character* (viz. his tendencies to lie, cheat, and steal). Accordingly, if he is to be an epistemically responsible agent, obeying standard epistemic norms, he should say to himself something like this: "Well, it has always seemed obvious to me that people have natural moral rights to life, liberty, and property – but now I have evidence that I believe this *because* I have tendencies to lie, cheat, and steal. I now also have empirical evidence that Rawls has a morally better personality than I. Whereas I tend to lie, cheat and steal, Rawls is honest, helpful, and kind. Given that Rawls is a better person than me, and given that Rawls *doesn't* think it is obvious at all that people have the natural rights I think they have, *I should give up my moral judgment in favor of Rawls'*."

Now, just to be clear, the toy case just discussed here is *not* intended to

be empirically realistic. I am not suggesting that the actual John Rawls was a good person and the actual Robert Nozick a worse person. The point of the example is instead a conceptual one about prevailing epistemic norms in moral and political philosophy – namely, that:

- *Epistemological Claim:* Prevailing epistemic norms in moral and political philosophy entail that we ought to aim, as far as possible, to epistemically privilege – in argument and theory-construction – moral premises endorsed by those among us who have the "best moral compass", as defined by reference to personality, character, and behavioral traits commonly recognized, by all parties to the relevant moral-political debates, to be morally good, bad, right, and wrong.

Clearly it is a further, complex empirical question whether it is possible to reliably determine who among us has the "best moral compass." The real world, after all, is not like our toy case. If our world contained people like our hypothetical "Rawls" and "Nozick", it would be easy to see who has the better moral compass – and thus, by common epistemic norms, easy to see whose moral and political premises we should epistemically privilege. The fact that the real world is not so simple, however, does nothing to cast doubt on *Epistemological* Claim – and I contend that we have just seen that *Epistemological Claim* is true: it *is* standard practice in moral and political philosophy to base arguments and theories on the moral premises that better people find attractive (i.e. "our premises") over premises found attractive by worse people (Nazis, psychopaths, and immoralists).

II. The Case for "Empirical Bet"

Suppose, then, that *Epistemological Claim* is true. Are there any good reasons to think empirical science *could* provide us with evidence of who among us have the "best moral compass"? This brings us to this paper's second primary claim:

- *Empirical Bet:* There *are* good reasons to believe that empirical philosophical-psychological research may be able to provide real (albeit imperfect) inductive evidence of (a) which individuals among us have the "best moral compass", and (b) which moral premises those people endorse.

What evidence is there in favor of *Empirical Bet*?

One obvious initial worry is that there may not be any, or at least not enough, personality, character, and behavioral traits that *all* parties to moral or political debates will agree to be good, bad, right, and wrong that might enable empirical scientists to determine who among us have the better (or best) moral psychology. Libertarians and liberal-egalitarians, for example, plausibly ascribe very different moral valences to very different psychological traits. Liberal-egalitarians, it seems, place pride-of-place on fairness, compassion, and kindness. They will say the fair, compassionate, and kind person is the good person. Libertarians, however, might think that liberal-egalitarians dramatically overestimate the moral value of these traits. Libertarians might say: "A little fairness, compassion, and kindness are okay – but too much of these things are morally terrible. It's more important to promote personal responsibility and self-efficacy than kindness, fairness, or compassion – because it's only personal responsibility and self-efficacy that make for a flourishing individual, and which, if instilled in the poor, will get them out of poverty. The liberal-egalitarian is *so* 'kind' that they just give handouts to the poor – which is really not good at all, because it provides the poor with perverse incentives to stay poor."[27] One can then, of course, imagine liberal-egalitarians responding with a very different story defending the moral superiority of *their* favored traits. This is one worry about *Empirical Bet*.

Another related worry is that it is actually impossible to specify "morally good" and "morally bad" psychological traits in a way that all parties to moral-political debates could accept. For example, one could imagine a libertarian saying to a liberal-egalitarian: "You *think* it's kind to give welfare to the poor – but it's not *kind* at all. It would be far more kind to get the poor to

[27] See e.g. A. Rand, *The Virtue of Selfishness*, New York, New American Library, 1964.

take responsibility for their lives. I know you will say this is callous, but *I* don't think it is. It's not callous; it's *kind*." In other words, the worry is that even to the extent that opposing parties can agree upon the moral valence of a given psychological trait or propensity – e.g. the goodness of kindness, the badness of callousness – there will never, or almost never, be ways of *specifying* the relevant traits that would satisfy the disagreeing parties.

A third, related worry is that even if all parties to a moral-political disagreement *did* agree upon the moral valence of a given trait (e.g. kindness is good), and all parties also agreed on a *specification* of kind behaviors (e.g. kindness is helping old ladies across streets, helping people truly in need until they can stand on their own two feet, etc.), it is hard, if not impossible, to see how these sorts of traits might be *operationalized* in empirical studies linking them to fundamental moral judgments of various sorts. What are we to do: see if self-described libertarians and liberal-egalitarians help old ladies across streets? Or how might we possibly operationalize, and study rigorously, "helping people truly in need"? Offhand, the empirical/methodological obstacles involved in measuring and studying who has a "better moral compass" can seem insurmountable.

Although these worries are worth taking seriously, there are reasons to think that they can all be surmounted. First, although there probably will be significant disagreement over which traits and behaviors are morally better than which – some may indeed claim that personal responsibility is morally better than compassion, whereas others may claim the opposite – it is an ultimately empirical question whether there is *enough* overlapping agreement about these issues to make it possible to arrive at a determinate, operationalizable, and widely acceptable account of which traits and behaviors define a "better moral compass." Second, although the objections discussed above aim to call the existence of such a consensus into question, there are good reasons to be optimistic. For although different people (e.g. coming from different philosophical perspectives) may have very different conceptions of *whether* certain traits or behavior (e.g. compassionate giving handouts to the poor) are morally good, or *how much* of a given trait or behavior is good (viz. giving

handouts is "too compassionate"), all parties to mainstream philosophical debates surely can agree on quite a great deal on many traits and behaviors as *clearly* reflecting a good or bad moral psychology. Every party to mainstream philosophical debates in moral and political philosophy, for example, will surely agree that general psychological propensities to lie, cheat, and steal, as well as propensities to *judge* that it is okay to do these sorts of things, are indicative of a bad moral psychology (or corrupt "moral compass"). A person who lies, cheats, and steals, and who thinks it is morally okay to do these things, is, *by all accounts*, a morally corrupt person.

This brings me to my primary response to the aforementioned worries, which is that *Epistemological Claim* – the claim that we ought to *aim* to epistemically prefer the moral judgments of those with the "best moral compass" – is entirely compatible with profoundly imperfect empirical evidence, including profoundly imperfect agreement on what constitutes a good or bad "moral compass." Here is why. Suppose there are many traits and behaviors that parties to mainstream philosophical debates disagree about the moral valence of. Again, to take an example mentioned earlier, suppose philosophers in one camp (e.g. libertarian political philosophers) tend to place a great deal of moral value on psychological traits supportive of *personal responsibility*, whereas philosophers in an opposing camp (e.g. liberal-egalitarians) tend to morally prioritize *compassionate giving* far above personal responsibility. Be that as it may, both camps may agree – and agree quite strongly – on some other indicators of, or what constitutes, a good or bad "moral compass." So, again, consider lying, cheating, and stealing, and the person who judges it is morally permissible to lie, cheat, or steal for their own benefit. It is not as though only some mainstream moral and political philosophers – say, liberal-egalitarians but not libertarians – would say that this sort of person has a profoundly corrupt moral compass. No, *all* mainstream moral and political philosophers – and certainly, the parties to the relevant debate (i.e. libertarians and liberal egalitarians) – would say this. Suppose then that empirical psychological research in fact found that individuals in one philosophical camp – e.g. people who find libertarian premises attractive – are

significantly more likely to lie, cheat, and steal, than individuals in the other camp. Even supposing that the two camps disagree about the moral valence of many other traits (e.g. compassion, personal responsibility), if it were found that individuals in one philosophical camp (libertarians) were significantly more likely to lie, cheat and steal than individuals in the other camp (liberal-egalitarians), that these tendencies to lie, cheat, and steal were *causally responsible* for libertarian moral judgments, and both camps *accept* that tendencies to lie, cheat, and steal are indicative of a corrupt moral compass (and again, both camps do accept this) it follows, per *Epistemological Claim*, that both camps would have *some real inductive evidence* that individuals in the one camp (libertarians) tend to have *seriously morally corrupt traits and propensities* that individuals on the other side don't have. Although this would only be one piece of evidence, it would nevertheless be *real and significant* empirical evidence that particular moral judgments about disputed premises – judgments favoring libertarian premises – are *in fact corrupted* by propensities to lie, cheat, and steal.

The general claim then is this: although there are all kinds of difficulties involved in obtaining agreement about which traits and propensities are morally better or worse (is compassion morally better than personal responsibility?), in operationalizing and measuring different traits (e.g. how does one operationalize kindness in order to study it empirically?), there are reasons to be optimistic that we can (A) achieve *some* significant agreement about the moral valence of particular traits and behaviors (e.g. all agree lying, cheating, and stealing for one's own benefit are *very* morally bad), and (B) successfully operationalize and study the relationships that *those* traits and behaviors bear to judgments about disputed moral premises. In other words, although our empirical evidence about what constitutes a good or bad "moral compass" may be *profoundly imperfect* – in that, again, there may be many traits we cannot agree about or effectively measure – we still may be able to obtain *strong empirical evidence* about traits, behaviors, and propensities that we *can* agree upon (e.g. lying, cheating, and stealing). And this is all that *Empirical Bet* requires. *Empirical Bet* reads that empirical philosophical-psychology can

provide "real (albeit imperfect) empirical evidence" of who among us, in philosophical debates, has the "best moral compass." Even if our evidence is very imperfect, we can still have *some* real inductive evidence that parties to one side of a philosophical debate endorse the premises they do as a result of certain corruptions of character. Although, again, this evidence might be very imperfect (and should be treated as such, epistemically), it would nevertheless be *real* empirical evidence – evidence that, *per Epistemological Claim*, should lead us to epistemically discount disputed premises linked to a corrupt moral personality.

One obvious worry at this point is that it is incredibly implausible to think that the kinds of relationships alluded to here – causal relationships between (clearly bad) propensities to *lie, cheat, and steal*, and judgments of disputed premises in moral and political philosophy – might really obtain in our world. Are we really to take seriously the idea that people who find, say, libertarian premises attractive in moral and political philosophy are more likely to *lie, cheat, and steal* than, say, liberal-egalitarians (or vice versa); that utilitarians might have greater propensities to lie, cheat or steal more than Kantians or virtue theorists; etc.? The very suggestion that this sort of thing might be the case might seem incredible.

The problem with this worry, however, is that it runs up against the grain of cutting-edge of empirical research. Recent findings in psychology and the neurosciences tentatively indicate that people who find certain moral and political premises attractive *may indeed* have greater tendencies to lie, cheat, and steal. For instance, several recent studies have found act-utilitarian judgments to be related to both psychopathy (tendencies to be cruel, etc.) and Machiavellianism (tendencies to deceive out of self-interest).[28] Economic libertarian judgments have been found to bear significant relationships to three particularly dark and anti-social traits: Machiavellianism (tendencies to deceive), Narcissism (overinflated self-worth and callousness), and Psychopathy (cruelty, and absence of guilt or remorse) – traits that have been

[28] A.L. Glenn, A. Raine, et R.A. Schug, « The neural correlates of moral decision-making in psychopathy », in *Molecular Psychiatry*, 14, 2009, pp.5-6; D.M. Bartels, et D. Pizarro, « The Mismeasure of Morals: Antisocial Personality Traits Predict Utilitarian Responses to Moral Dilemmas », in *Cognition*, 121, 2011, pp.154-161; M. Arvan, « A Lot More Bad News for Conservatives, and a Little Bit of Bad News for Liberals? Moral Judgments and the Dark Triad Personality Traits: A Follow-up Study », *Neuroethics*, 6/1, 2013, pp.51-64.

shown to be strongly related to immoral *overt* behavior[29], and which are intuitively related to or comprise traits that ordinary people and philosophers widely consider to be moral vices (e.g. cruelty, insensitivity, lack of remorse for wrongdoing, etc.).[30] Similarly, socially conservative views on moral issues – issues ranging from gay marriage to capital punishment – have found to be systematically related to the same three dark traits across a wide array of issues (to give you an idea of just how systematic these relationships were, across my two studies, I found *1.54* significant relationships between socially conservative moral judgments and the three members of the Dark Triad per moral issue examined, compared to only *0.06* relationships for socially liberal judgments [e.g. judgments favoring the permissibility of gay marriage, against capital punishment, etc.]).[31] These results are not only provocative; they strongly suggest that *Empirical Bet has some evidence in its favor*. There are reasons to think that empirical philosophical-psychology may well discover that some moral and political judgments about fundamental premises – judgments favoring libertarianism, social conservativism, and act-utilitarianism – are strongly related to morally corrupt psychological propensities and patterns of behavior. At the very least, there are reasons to be optimistic. Thus, as many difficulties as there may indeed be fixing in on and measuring traits, propensities, and behaviors that philosophers in disputing camps may agree upon as indicative of a good or bad "moral compass", etc., there are reasons to be optimistic about *Empirical Bet.*

[29] D.N. Jones, D.N., et D.L. Paulhus, D.L., « Differentiating the Dark Triad within the interpersonal circumplex », in *Handbook of interpersonal theory and research*, New York, Guilford, 2010, pp. 249-267; D.N. Jones, et D.L. Paulhus, « Different Provocations Trigger Aggression in Narcissists and Psychopaths », in *Social and Personality Psychology Science*, 1, 2010, pp. 12-18; D.N. Jones, et D.L. Paulhus, « The role of impulsivity in the Dark Triad of personality », *Personality and Individual Differences*, 51, 2011, pp. 679-682; and Nathanson et al. (2006).

[30] M. Arvan, « Bad News for Conservatives? Moral Judgments and the Dark Triad Personality Traits », in *Neuroethics*, 6/2, 2013, pp.307-318.

[31] M. Arvan, « A Lot More Bad News for Conservatives, and a Little Bit of Bad News for Liberals? Moral Judgments and the Dark Triad Personality Traits: A Follow-up Study », in *Neuroethics*, 6/1, 2013, pp.51-64.

III. Objections, and Replies[32]

Objection #1 – One doubt about Empirical Bet (the "Skepticism-About-Character-Traits Objection"): "Some people (e.g. Doris, 2002, Harman, 2000) argue that empirical data show that there are no genuine character traits, or, if there are, such traits do not play the role we think they play in moral behaviour. If these critiques are correct, this would constitute a fatal objection to your 'experimental ethics'."

Reply: Indeed, such arguments have been given. However, there also appear to be compelling grounds for resisting them. The standard response is that skeptics about character traits (e.g. Doris and Harman) mischaracterize the nature of character traits and the role such traits play in morality and human behavior. Doris and Harman appear to have a crude behaviorist conception of character traits. They argue that (e.g.) there is no such thing as courage because whether or not people do courageous things depends profoundly on situational aspects outside of the person. Doris and Harman tacitly assume, then, that character traits must be stable dispositions to behave in certain ways across a wide variety of situations. Their argument is that because different situations have profound effects on what people do, there are no such traits. But, while the empirical data do call the existence of these kinds of traits into question – that there is no such thing as courage in that sense, if people are so susceptible to environmental/situational effects – this is not how character traits are understood in traditional Aristotelian virtue theories of morality. Aristotelians understand character traits as situation specific – as dispositions to behave in certain ways in response to specific kinds of situations and stimuli.[33] Yet, far from disproving the existence of these kinds of situation-specific dispositions,

[32] I would like to thank two anonymous referees for pressing the objections raised in this section. For the sake of addressing their concerns as raised, I have chosen to quote their objections directly, with some minor editorial alterations.

[33] See e.g. J. Annas, *The Morality of Happiness*, Oxford, Oxford University Press, 1995; R. Kamtekar, « Situationism and Virtue Ethics on the Content of Our Character », in *Ethics*, 114/3, 2004, pp. 458-491; R.C. Solomon, « What's Character Got to Do With It? », in *Philosophy and Phenomenological Research*, 71/3, 2005, pp. 648-655; and J. Webber, « Virtue, Character, and Situation », in *Journal of Moral Philosophy*, 3/2, pp. 193-213. Also see S.M. Samuels, M. Steven and W.D. Casebeer, « A Social Psychological View of Morality: Why Knowledge of Situational Influences on Behavior Can Improve Character Development Practices », in *Journal of Moral Education*, 34/1, 2005, pp. 73-87, who argue that empirical findings at most suggest we should alter practices of moral education to be more sensitive to situations.

the empirical evidence that people like Doris and Harman appeal to actually appears to confirm the existence of such traits. For example, Doris himself refers to a series of studies on student behavior which indicate that although there is no trait such as "student honesty" that spans all situations – for example, the probability of whether a student is likely to cheat on an exam has found to be unrelated to the probability of the child stealing unattended money or faking records of athletic performance – students do display consistent dispositions to be honest or dishonest ways in *specific* situations: for instance, exams.[34] Students who cheat (or do not cheat) on one exam, for instance, are likely to behave the very same way on exams in other situations. Thus, while the empirical evidence does suggest that there are no "unified" character traits such as honesty per se, it also suggests that there are situation-specific traits such as exam-honesty, athletic-honesty, unattended-money-honesty, etc.

We cannot resolve these issues here. Whether there are robust (psycho-behavioral) character traits – and what those traits are like – are *empirical* questions not yet fully answered. Accordingly, it is perfectly legitimate to consider it an *open question* whether there are such traits – and this is all the present paper assumes. All this paper is make an Empirical Bet: that there *are* such traits, and that some such traits bear relevant relationships to judgments about moral and political premises. And given the empirical evidence above, it seems clearly premature decide this bet one way or the other at the present time. It is a legitimate bet to make at the present time.

Objection #2 – A second doubt about Empirical Bet (the "No-Way-to-Detect-Causal-Effects-of-Character-Traits-on-Moral-Political-Judgments Objection")*:* "I just don't see how it can be shown empirically that tendencies to cheat or lie could by themselves be *causally responsible* for (say) libertarian moral judgements."

Reply: Empirical Bet does not assert that tendencies to behave badly (e.g. to cheat or lie) are by themselves causally responsible for specific moral

[34] J. Doris, *Lack of Character: Personality and Moral Behavior*, Cambridge, Cambridge University Press, 2005, p. 63. For the studies themselves, see H. Hartshorne and M.A. May, *Studies in the nature of character: Vol. 1. Studies in deceit*, New York, Macmillan, 1928, pp. 379-80, p. 411. Also see J. Webber, « Virtue, Character, and Situation », in *Journal of Moral Philosophy*, 3/2, pp. 193-213.

judgments, only that they can be a *part* of what causes a person to make the relevant judgments, such that (*qua Epistemological Claim*) we should take the judgments to be corrupted by morally bad biases. The relevant question, in other words, is not whether dispositions to lie or cheat are the sole causes for people making particular moral or political judgments; the question is whether the dispositions can be discovered to play *a significant* causal role in biasing such judgments. And there is a good amount of preliminary evidence that this is likely the case. I have found, for instance, that people more highly disposed to deceive others (e.g. people who score higher on the Machiavellian personality trait) are more likely to make libertarian economic judgments over people who are less disposed to deceive.[35] Similar, I along with several others have found relationships between certain utilitarian judgments in trolley cases that a vast majority of ordinary people and philosophers consider to be plainly immoral (e.g. pushing a fat man in front of a trolley to save five lives) and psychopathic and Machiavellian personality traits.[36] Finally, for the sake of the present paper, I completed a simple study of 200 participants (recruited online at Amazon Mechanical Turk for $0.50 compensation per participant) examining relationships between the Dark Triad personality traits (Machiavellianism, Narcissism, and Psychopathy), as measured by the Short D3[37] – character traits that have been found to be *causally* related to immoral behavior[38] – and the following list of basic moral judgments (which subjects rated on a standard 1-5 Likert scale, with 1 indicating "strongly disagree" and 5 indicating "strongly agree"):

---

[35] M. Arvan, « Bad News for Conservatives? Moral Judgments and the Dark Triad Personality Traits », in *Neuroethics*, 6/2, 2013, pp. 307-318.

[36] *Ibid*. Also see A.L. Glenn, A. Raine and R.A. Schug, « The neural correlates of moral decision-making in psychopathy », in *Molecular Psychiatry*, 14, 2009, pp.5-6; D.M. Bartels, et D. Pizarro, « The Mismeasure of Morals: Antisocial Personality Traits Predict Utilitarian Responses to Moral Dilemmas », in *Cognition*, 121, 2011, pp. 154-161.

[37] D.N. Jones, D.N., et D.L. Paulhus, D.L., « Differentiating the Dark Triad within the interpersonal circumplex », in *Handbook of interpersonal theory and research*, New York, Guilford, 2010, pp. 249-267; M. Arvan, « Bad News for Conservatives? Moral Judgments and the Dark Triad Personality Traits », in *Neuroethics*, 6/2, 2013, pp. 307-318.

[38] See D.N. Jones, D.N., et D.L. Paulhus, D.L., « Differentiating the Dark Triad within the interpersonal circumplex », in *Handbook of interpersonal theory and research*, New York, Guilford, 2010, pp. 249-267; D.N. Jones, et D.L. Paulhus, « Different Provocations Trigger Aggression in Narcissists and Psychopaths », in *Social and Personality Psychology Science*, 1, 2010, pp. 12-18; D.N. Jones, et D.L. Paulhus, « The role of impulsivity in the Dark Triad of personality », *Personality and Individual Differences*, 51, 2011, pp. 679-682.

"It is wrong to kill"
"It is wrong to lie"
"It is wrong to steal"
"It is not that bad to lie"

The results of the study are striking (see Table A). Psychopathy was significantly correlated with disagreement with the moral judgments that it is wrong to kill and steal, and bad to lie. Further, all three traits were strongly related to agreement with the moral judgment that it is "not that bad" to lie.

	Machiavellianism	**Narcissism**	**Psychopathy**
"It is wrong to kill"	r = .035, p = .625	r = -.012, p = .869	*r = -.189, p = .007***
"It is wrong to lie"	r = -.131, p = .066	r = -.005, p = .948	r = -.126, p = .126
"It is wrong to steal"	r = .026, p = .714	r = -.038, p = .592	*r = -.262, p<.00001***
"It is not that bad to lie"	*r = .268, p < .00001***	*r = .152, p = .032**	*r = .287, p = <.00001***

*Correlation is significant at .05 level (2-tailed).
**Correlation is significant at .01 level (2-tailed).

Table A.

Although these studies only examine correlations – and the exact causal relationships between moral judgments and psycho-behavioral traits are unsettled – there are emerging lines of research that strongly suggesting that psycho-behavioral traits do play a causal role in producing different moral and political judgments.[39]

[39] See e.g. J.A. Terrizzi Jr., N.J. Shook and L. Ventis, « Disgust: A predictor of social conservatism and prejudicial attitudes toward homosexuals », in *Personality and Individual Differences*, 49/6, 2010, pp. 587-592; H. Thórisdóttir and J.T. Jost, « Motivated ClosedMindedness Mediates the Effect of Threat on Political Conservatism », in *Political Psychology*, 32/5, 2011, pp. 785-811.

Accordingly, although *Empirical Bet* is by no means settled, there are plenty of reasons to think it likely that the bet will be borne out by further empirical study – which is all this paper assumes. Cutting-edge empirical research suggests that there probably are causal relationships from psycho-behavioral traits to moral and political judgments, and that further investigation should be able to determine what those causal relationships are. After all, such causal relationships – if they exist – should manifest themselves neurologically, and probably behaviorally as well. Here are just a few ways that such causal relationships might be detected: First, neuroscientists should presumably be able to pin down, in the foreseeable future, areas of the human brains causally responsible for particular moral judgments. After all, neuroscientists have already isolated areas of the brain that represent particular memories, visual-spatial patterns, etc. Insofar as moral and political judgments are presumably represented by the human brain, it should be possible for cognitive neuroscientists to pin down where (and how) those judgments are represented in the brain. Next, it should be possible to determine whether those representations are causally "upstream" or "downstream" to areas of the brain causally responsible for behavioral tendencies (e.g. tendencies to lie, cheat, steal, etc.). Accordingly, there is every reason to believe that whatever causal relationships exist between psycho-behavioral traits and moral-political judgments can be detected by empirical science.

Here are two studies I hope to carry out myself in the near future. In the first study, participants in a virtual reality (VR) setting – an experimental setting increasingly being used in empirical psychology to study human behavior and cognition – might be primed to display morally bad tendencies. That is, they might be instructed to commit wanton acts of theft, vandalism, lying, even murder, within a virtual reality simulation. Immediately after so primed, experimenters might present subjects with moral-political judgment surveys, such as Arvan's MIS surveys (which ask respondents to make moral-political judgments on an array of issues).[40] Because changes in dependent variables in response to variations in independent variables indicate causal

[40] See M. Arvan, « Bad News for Conservatives? Moral Judgments and the Dark Triad Personality Traits », in *Neuroethics*, 6/2, 2013, pp.307-318; and M. Arvan, « A Lot More Bad News for Conservatives, and a Little Bit of Bad News for Liberals? Moral Judgments and the Dark Triad Personality Traits: A Follow-up Study », Neuroethics, 6/1, 2013, p. 63.

dependence of the dependent on independent variables (the method of investigating causal dependence known as concomitant variation), any effects on subjects' moral judgments caused by priming behavioral tendencies would strongly indicate causal dependence of moral-political judgments on the relevant behavioral traits. Similarly, we might test causal dependence in the opposite direction – to determine whether variations of moral-political judgments cause differences in behavioral tendencies – by priming subjects to make certain moral judgments before entering a virtual reality setting (e.g. by presenting subjects with a series of moral and political assertions, such as "It is morally right to tax citizens to help the poor", etc.) and then seeing whether priming them with those judgments affects the way they behave in the simulation (e.g. lying, cheating, stealing, etc.). These are just some possible ways to examine causal relationships between behavioral tendencies and moral-political judgments. Surely there are other possible ways as well. *If* there are such causal relationships – as *Empirical Bet* wagers – there is every reason to believe that they should be detectable.

Objection #3 – A third doubt about Empirical Bet (the "No-*Necessary*-Causal-Relationship-Between-Character-Traits-and-Moral-Political-Judgments Objection"*)*: "Your argument presupposes that if someone is disposed to do morally nasty things, they *must* also be bad at judging what is right and what is wrong – for instance, that someone who steals *must*, at some level at least, also think that stealing is okay, or not that bad, or at any rate, must value honesty less than someone who never steals. Yet, although it could be the case that one person's vicious character leads them to make 'bad' moral judgements, it doesn't seem to be necessarily the case for all persons with vicious character. Further, even if it turned out that there is a reliable connection between vice and moral judgement (in a similar way there may be a reliable connection between virtue and moral judgement), it isn't clear that morally bad behaviour is always a product of vicious dispositions […] some morally problematic behaviours are disconnected from the person's evaluative outlook (think of the obedient subjects in the Milgram experiments—do we want to say that they are

incapable of making justified moral judgments?). Finally, although vices may corrupt day-to-day moral judgements, there doesn't seem to be reasons to think that they should corrupt the highly theoretical moral judgements that we make in moral debates. Even though a Nazi may (wrongly) think that the Aryan race should have supremacy, he could still be a Nozickian, a Rawlsian, a utilitarian, and even a virtue ethicist, when he is debating about morality and justice with his friends at the pub."

Reply: Empirical Bet doesn't presuppose that a person who is disposed to do nasty things *must* be bad at judging right from wrong, or that morally bad behavior is *always* a product of vicious dispositions (surely, as Doris, Harman, and many others have shown, situations can provoke bad behavior, too). *Empirical Bet* only presupposes that there are *some* significant causal relationships between psycho-behavioral traits and moral-political judgments that can *bias and improve* those judgments, the empirical detection of which would thereby (*qua Epistemological Claim*) give us good grounds to some judgments over others in moral-political argument and theory construction. Second, the worry that while vices may corrupt ordinary, everyday moral judgments but not corrupt highly theoretical judgments is belied by a growing body of cutting-edge empirical work linking morally dubious traits and behaviors to *specific* highly theoretical judgments (see e.g. my research linking all three traits of the Dark Triad to libertarian judgments, etc.).[41] Finally, as we have already seen, one of the most standard epistemological methodologies in moral and political philosophy is to make arguments, and evaluate highly theoretical principles (and theories themselves), on the basis of moral judgments about *cases*. So, for example, utilitarianism is commonly criticized as a moral theory on the basis of its implications in Organ Donor case discussed earlier, as well as in Trolley Cases (it implies, for instance, that one ought to push a fat man in front of a Trolley to save multiple lives, at the cost of the fat man's life – an implication that a vast majority of people judge to be immoral[42]). Since a growing body of research suggests judgments about

[41] Again, see M. Arvan, « Bad News for Conservatives? Moral Judgments and the Dark Triad Personality Traits », in *Neuroethics*, 6/2, 2013, pp.307-318.

[42] See J.J. Thompson, « Killing, Letting Die, and the Trolley Problem », in *The Monist*, 59/2, 1976, pp. 204-217; J.J Thompson, « Turning the Trolley », in *Philosophy and Public Affairs*, 36/4, pp. 359-374; and P. Singer, « Ethics and Intuitions », in *Journal of Ethics*, 9/3-4, 2005, pp. 331-352.

specific *cases* may be corrupted by bad psycho-behavioral traits – utilitarian judgments in the "fat man" Trolley Case have been repeatedly linked to Machiavellianism and Psychopathy[43] – there are reasons to believe (*qua Empirical Bet*) that empirical research on judgments about cases should have implications (*qua Epistemological Claim*) for how we evaluate high-level principles and theories.

Objection #4 – One doubt about the argument for Epistemological Claim (the "We-Don't-Evaluate-Principles-on-the-Basis-of-Character Objection"): "Your argument appealed to how we standardly evaluate a Nazi's premise that the Aryan race ought to dominate the world. You implied it is standard practice to reject the premise on the basis of the Nazi's character. But this seems incorrect. We reject the Nazi's premise not on grounds of the Nazi's character, but rather on substantive grounds – that is, by appeal to substantive moral argument based on moral principles (e.g. the principles that discriminating against people on the basis of race is arbitrary, and morality is non-arbitrary). What justifies our ignoring Hitler's moral beliefs in theory construction is that his beliefs are *substantively wrong*, not that he is evil."

Reply: On the contrary, I submit it is evident upon further reflection that many (if not all) so-called "substantive arguments based on moral principles" are really, at a deeper level, epistemic appeals based on judgments about moral character (in line with *Epistemological Claim*). Allow me to use two well-known examples to illustrate.

Consider first the most infamous objection to classical act-utilitarianism: the objection that it fails to properly account for *justice* or *moral rights*. This objection is often motivated by reference to examples. One example is commonly known as the Organ Donor Case[44]: a doctor on a

[43] Again, see A.L. Glenn, A. Raine and R.A. Schug, « The neural correlates of moral decision-making in psychopathy », in *Molecular Psychiatry*, 14, 2009, pp.5-6; D.M. Bartels, et D. Pizarro, « The Mismeasure of Morals: Antisocial Personality Traits Predict Utilitarian Responses to Moral Dilemmas », in Cognition, 121, 2011, pp.154-161; M. Arvan, « Bad News for Conservatives? Moral Judgments and the Dark Triad Personality Traits », in Neuroethics, 6/2, 2013, pp.307-318.

[44] See e.g. M. Timmons, *Moral Theory: An Introduction, 2nd ed.*, Lanham, Rowman and Littlefield Publishers, 2013, pp. 144-145.

transplant-ward has five patients who will die if not given new organ transplants immediately. Further, all five patients are not only well-loved (they have numerous family members and friends who care deeply about them); they are also socially important: one is, say, a president of a large corporation, one is a famous musician; etc. Then, on the other hand, the doctor has a lonely, miserable bum off the street walk in for treatment for a common cold. Finally, suppose that the doctor *knows* that no one will miss the bum if he dies, and that there is no realistic way to obtain organs necessary for saving the other five patients aside from (covertly) killing the bum to harvest the bum's organ's. Offhand, act-utilitarianism plainly entails that it would be morally right for the doctor to kill the bum for his organs, as doing so would (of all the available options available) maximize happiness (no one will find out; the bum won't be missed; five well-loved and socially important lives will be saved; etc.). But, the objection goes, this is substantively wrong. Any moral theory which entails that it is permissible (not to mention right!) to kill healthy innocent persons to harvest their organs for the good of others is a theory so out of line with our other moral commitments that we cannot accept it.

This looks, at least at first glance, like a substantive moral argument against act-utilitarianism – one based not on judgments about a person's character but instead on grounds of moral principle. But is it really a substantive argument *at bottom*? It is hard to see how it is. After all, it is not as though there aren't very compelling-looking principled arguments *for* act-utilitarianism. The usual argument for act-utilitarianism looks compelling enough: (1) Each person's happiness is their good, (2) Moral rightness requires doing what's best (i.e. maximizing the good), so, (3) Moral rightness requires maximizing happiness (in every action). On what grounds, then, is the "justice" objection against utilitarianism made? It is not, after all, as though *all* philosophical inquirers accept the objection. No, there are utilitarians out there who respond to the justice objection though *bold denial*: claiming that, whatever our intuitions or feelings about Organ Donor-type cases may be, it is strictly *right* to kill one person and harvest their organs to save multiple lives.[45]

[45] See e.g. M. Timmons, *Moral Theory: An Introduction, 2nd ed.*, Lanham, Rowman and Littlefield Publishers, 2013, p. 150.

But this bold denial strategy is rarely taken very seriously. Why? Part of the reason, surely, is that there are other, alternative theories – and substantive moral principles – out there that intuitively appeal to us more: for instance, Kantianism, which requires us to always treat people as ends-in-themselves, not mere means to be used for the good of others. But is this the *whole* story, or even, *at bottom*, the correct story? For notice that the next question which inevitably arises is this: *why* do we find, say, Kantianism "more compelling" in the relevant regard? The answer, it seems, is this: *we think that any morally decent person has the Kantian intuition that morality doesn't permit people to be treated as mere means (or objects), not the act-utilitarian intuition*. We rest our argument, in other words, on claims about moral character. We (mostly) ignore act-utilitarians who engage in the strategy of bold-denial *because we think that no morally decent person can take the bold denial seriously*.

The same appeal to moral character also underlies, I think, one of the most notorious objections to Kant's moral theory (at least as Kant understood it). Kant, of course, seems to argue that it is *never* right to lie, even if there is a murderer at your door looking to kill an innocent person and you might save the innocent person's life by lying.[46] This too strikes almost everyone who does moral philosophy as patently absurd. If that's what Kant's theory entails, so much the worse for his theory. But in that case the same question arises as in the utilitarian case: *why* do we object to this supposed implication of Kant's theory? Again, I think the most compelling answer is this: we think that *no decent person* could accept the implication. Any decent person, we think, would consider lying *right* (and at least permissible) in order to save innocent lives.

Now, the objector I am facing might try to deny that we ultimately appeal to character in these cases. They might say: "The objections are based on *appeals to intuition* or *substantive moral principles*, not character" – to which I say, again: yes, but only the intuitions and substantive moral principles of *some people:* morally decent people "like you" – not the intuitions or moral principles of psychopaths, immoralists, or Nazis. Indeed, this is the crucial point: *not* everyone has the relevant intuitions or accepts the "correct

[46] I. Kant, « On a supposed right to lie from philanthropy », in *Immanuel Kant Practical Philosophy*, Cambridge, Cambridge University Press, 1996, pp. 611-615.

substantive moral principle." Psychopaths, for example, see no reason not to lie *at all*, not to mention in cases like Kant's murderer-at-the-door. And again, there are *some* utilitarians out there who are prepared to say it *is* right to kill an innocent person for their organs. How do we argue against these people? Not, it seems, on grounds of substantive moral principle (they *reject* the principles we appeal to). No, we argue against them by resting on appeals to what "we" – morally decent people in general – find "intuitive." The problem of justice for utilitarianism is a *problem* for utilitarianism, and the problem of lying to the murderer a *problem* for Kant's construal of his moral theory, simply because *we, decent people, judge it that way*.

If there is any remaining doubt about this, I encourage readers to engage in a line of thought suggested by Judith Lichtenberg in her paper, "Moral Certainty."[47] Lichtenberg asks us to reflect on what we are more certain about in ethics: *substantive moral principles* or *particular moral experiences*. So, for example, Lichtenberg asks us: imagine seeing a young child being tortured on the street-corner. Any decent person, Lichtenberg suggests, will immediately experience the action *as wrong*. It is what we are most certain about. Any moral theory or principle which conflicted with *that* moral experience would be less certain – as a moral theory or principle – than our direct experience of the action's wrongness. But again, it cannot *merely* be our certainty that guides us. Not everyone experiences a child's torturing the same way: the psychopath certainly sees nothing wrong with it. No, we see it *as* wrong – and we hold up moral principles and theories *against* our experiences as test-cases – *qua morally better people*. Further, when people deviate from such judgments (e.g. when they see "nothing wrong" with things the rest of us take to be wrong), we typically *question the person's motives*[48] *and/or character*:

> [I]f someone fails to see what is wrong […] with Jeffrey Dahmer, who killed and mutilated people, collected their bones and ate their flesh, we do not conclude simply that their 'values' differ from ours, but that something has gone seriously wrong with them. Their problem rests not on a mistake but in an affective, emotional defect.[49]

[47] J. Lichtenberg, « Moral certainty », in *Philosophy*, 69, 1994, pp. 181-204.
[48] *Ibid*., p. 186.
[49] *Ibid*., p. 187.

In short, although our reasons for rejecting particular moral or political principles (e.g. the Nazi's principle that Aryans should rule the world) are *superficially* based on substantive moral argument, I contend that it is clear, at a deeper level, that most (if not all) such arguments *ultimately* trace back to judgments about character: namely, that the principles *we* affirm – about not killing innocent people for their organs, etc. – are true or justified *because people-like-us, people with decent moral character, accept them.*

Objection #5 – A second doubt about the argument for Epistemological Claim (the "You've-Overstated-Standard-Epistemological-Practice Objection")*:* "You claim it is standard practice in moral and political philosophy to ascribe greater credence to moral judgments endorsed by *virtuous* agents over less virtuous or vicious ones – but the evidence you offer suggests a different philosophical practice and a much weaker epistemic principle: namely, the practice/principle of basing moral and political arguments on premises that *morally decent, ordinary people* are apt to accept."

Reply: I agree that it currently is standard practice to base moral and political arguments on this weaker epistemic standard (i.e. the moral principles morally decent people like us find attractive are apt to accept). *Epistemological Claim* holds, however, that this practice entails that we *ought* to adopt a higher epistemic standard: basing moral and political arguments on premises *the most virtuous* among us are apt to accept. My claim, in other words, is not that standard practice in moral and political theory *actually embodies* this higher standard, but rather that it *entails* such a standard. And I think the grounds for this entailment are clear. One cannot make the first, weaker epistemic move (viz. "Moral and political arguments should be based on premises *ordinary, decent people* find attractive"), I submit, without presupposing grounds in favour of *Epistemological Claim* ("Moral and political arguments should be based on premises the best people among us are apt to accept"). Here's why: if we are looking for truth – which I assume we are doing in philosophy – any argument for the weaker epistemic principle (i.e. of basing moral and political

arguments on premises decent people accept) has to have an *epistemic basis*. But what epistemic bases could there be for privileging the premises decent people find attractive over, say, the premises the psychopath is apt to accept? One has to say something like this: "We should base moral and political arguments on the premises ordinary, decent people are apt to accept over the psychopath, because ordinary, decent people *have a better moral sense*. Decent people are *less-biased* on grounds we all consider to be immoral – they're not tempted to lie, cheat, or steal like the psychopath." My claim is that once we recognize that decent people have a "better moral compass" than the psychopath (and we think it is *therefore* epistemically better to base arguments on the decent person's premises than the psychopaths), by parity of reasoning there is simply *no avoiding* the stronger conclusion that it would be epistemically better to base moral and political arguments and theories on the premises *the best, most virtuous* people among us are apt to accept over premises that appeal to merely decent people.

Objection #6 – A third doubt about Epistemological Claim (The "Your-Epistemology-Undermines-Your-Overall-Proposal Objection")*:* "The fact that 70% (or even 80, or even 100%) of psychopaths make utilitarian judgements (or whatever), although surprising, does not make me believe that the typical utilitarian's premises are false. After all, many people who are *not* psychopaths are utilitarians […] The problem then is that arguably there are virtuous and unvirtuous people on *all* sides of major debates. But since, according to *Epistemological Claim*, we should privilege the moral-political judgments of *good* people, it follows that empirical research can't help us at all: there are good people who find Kantianism attractive, good people who find Utilitarianism attractive, good people who find Virtue Theory attractive, etc. There is perhaps a required threshold of *moral decency*, but it seems that most people engaged in the relevant debates may satisfy it."

Reply: Suppose one *is* a virtuous person, but in terms of some small subset of one's moral or political beliefs – one is a staunch libertarian, let's suppose – one "finds oneself in bad company" (i.e. it turns out, empirically, that

some large percentage of people who find libertarian premises attractive have morally bad behavioral tendencies). What is the proper epistemic response for you, an otherwise good person, to take in this situation? I propose, *contra* the objection, that there is a very strong epistemic case for the proposition that "numbers" – the sheer number of bad people who find a given moral-political judgment attractive versus the number of good people – should affect how you judge *the overall balance of evidence* in favor of or against the truth of the moral-political judgments you share with those bad people. Indeed, the following seems to me to be clearly the most epistemically responsible way to respond to the situation. One should say to oneself: "Evidently, there is *something* about the [libertarian] moral judgment I favor that, for some reason or other, attracts very bad people to it. This suggests – even though I can't seem to see it (perhaps my other virtues *blind* me to it) – there is some "bad-attracting" feature of the judgment(s) in question, i.e. morally bad properties. The fact that *I* find the judgment attractive too – as a good person – may suggest that the judgment also has *some* morally good properties as well – but given that a vast majority of other virtuous people (people who, like me, behave well and treat others in morally good ways) *deny* the truth of my libertarian judgment, the epistemically responsible thing for me to conclude is that even if the judgment has *some* morally good properties which presumably attract me to it, it has an even *larger* balance of morally bad properties that attract many bad people and repel most other good people."

Why, exactly, do I think this is the most epistemically reasonable way to respond? Again, I submit that it is *standard epistemic practice* in moral and political philosophy. Go back, for example, to the two cases I discussed earlier: to the Organ-Donor counterexample to Utilitarianism and the Lying-to-the-Murderer-at-the-Door counterexample to Kant's absolutist judgment that lying is always wrong. It is perhaps *possible* for a good person to truly believe that it is right to kill an innocent bum to give their organs to five people (after all, it *would* save five lives!), or for a good person to truly believe that it is never right to lie (all indications are that Kant himself was a pretty decent chap). Be that as I may, I do not know of many moral or political philosophers who think we

should assign much *epistemic credence* to either view on such grounds. It is standard practice in moral philosophy to simply reject those beliefs in argument and theory construction – even though *some* good people might have them – precisely because the *vast* majority of good people who consider the cases agree that the judgments are wrong. Accordingly, the objection only holds if prevailing epistemic standards in moral and political philosophy – of epistemically privileging the premises accepted by *vast* majorities of good people (i.e. it is wrong to kill innocent people for their organs) over the premises found attractive by only a few good people on the fringes (i.e. it is okay to kill innocent people for their organs if doing so will maximize happiness) – are *unjustified.* Although, again, there may indeed be some who wish to hold that these standards *are* unjustified, those who think this should take it up with the profession, not the present paper. It is surely worthwhile to investigate *what prevailing epistemic standards imply*, which is all the present paper aims to do (again, see *Epistemological Claim*). Whether prevailing epistemic standards in moral and political philosophy are truly justified is a very broad issue warranting (at least) another full article in its own right. We cannot settle such issues here.

Objection #7 – A fourth doubt about Epistemological Claim (The "What-Really-Matters-Epistemically-is-Whether-*A*-Good-Person-Could-Give-Good-*Substantive*-Reasons-for-the-Premise, Not-How-Many-Good-People-Find-it-Attractive Objection")*:* "The 'Nozick' vs. 'Rawls' thought experiment actually looks like a *reductio* of the proposed procedure than an argument for it. Given that Nozick's premise *is* reasonable, and could be/is accepted by decent people, finding out that 'Nozick' has vices should not make us disregard it…At best, knowing that a given philosopher is not a very good person should make us consider whether the views he or she is endorsing could be defended by someone who *is* a good person. An alternative procedure, one which I suspect is already in currency, would therefore be to ask people to give reasons for their premises. Overall, it looks like the author is attempting to legitimize the use of 'genetic fallacies' in the context of moral/political debate, a practice that to this

day still finds no reason to be deemed acceptable in any serious philosophical inquiry."

Reply: The objection just brings us back to the problem this paper began with. People disagree over moral and political premises. Arguments "bottom out." The question we began with is this: when arguments bottom out in such a way – when two or more good or decent people both find fundamentally opposing attractive, and no further substantive argument can be given to break the deadlock – are there really no further epistemic grounds, grounds that *all* parties to the relevant debates can accept as true, to adjudicate the debate *in favour* of one side's premises over the other? I have argued here that there may indeed be (contingent upon the results of **Empirical Bet**). First, we all agree that certain behaviors and dispositions are generally right and good – honesty, conscientiousness, compassion – and other behaviors and dispositions (lying, stealing, and killing) bad and wrong. Second, there are some reasons to think that our judgments about moral-political premises can be improved by character traits that dispose us to behave in good ways, and corrupted by character traits that dispose us to behave in bad ways (*qua* **Epistemological Claim**). My argument has been that *if* these two things are the case – if **Empirical Bet** pans out – we will have strong *epistemic* reasons to do more than simply ask people to give reasons for their premises: we will have epistemically justified grounds for engaging in the *empirically informed moral epistemology* this paper proposes.

IV. Conclusion

This paper has defended the following two claims:

- *Epistemological Claim***:** Prevailing epistemological norms in moral and political philosophy entail that we ought to aim, as far as possible, to epistemically privilege – in argument and theory-construction – moral premises endorsed by those among us who have the "best moral compass", as defined by reference to personality, character, and behavioral traits commonly recognized, by all parties to the relevant

moral-political debates, to be morally good, bad, right, and wrong.

- *Empirical Bet:* There are good reasons to believe that empirical philosophical-psychological research *may* be able to provide real (albeit imperfect) inductive evidence of (a) which individuals among us have the "best moral compass", and (b) which moral premises those people endorse.

Accordingly, if this paper has been successful, philosophers and psychologists have compelling epistemological reasons to engage in a vast new research program: an *experimental ethics* – an *empirically-informed moral-virtue epistemology* – that aims to utilize empirical research on personality, character, and overt behavior to inform our judgments about the types of premises it is legitimate to invoke in moral and political philosophy.