4.10.18 | **Jeremy Fischer**

Response

Method in Moral Psychology

"Moral psychology" is a recently-invented term of art with a somewhat disputed reference, which reflects disagreement about its proper method(s).1 Some use the term to designate venerable philosophical inquiry into the nature and moral significance of psychological states. These inquiries are carried out through armchair reflection based in common sense, everyday observation, or intuition about, for instance, what moral responsibility is and whether it requires volitional activity; what emotions are and when, if ever, they are appropriate; and (the old holy grail of ethics) what happiness is and whether ethical virtue always delivers it. The Socratic paradoxes—such as that virtue is knowledge, that weakness of the will is impossible, that a good person cannot be harmed —comprise perhaps the earliest recorded instances of this style of moral psychology in European philosophy.

0
Shares

f

y

♥

Some recent practitioners—dissatisfied with the impoverished results of what we might call *reflective moral psychology*—have instead carried out their inquiries largely through humanistic study, using historical, literary, and sociological results and the interpretative methods of those

disciplines. We may call this *humanistic moral psychology*. In this category we find Bernard Williams's use of Sophocles's *Ajax* to provide an account of shame and moral incapacity,2 Martha Nussbaum's interpretation of Aeschylus's *Oresteia* in service of her accounts of anger and forgiveness,3 Gabriele Taylor's analysis of Shakespeare's *Coriolanus* in her monograph on vice,4 and Charles Mills's study of John Hearne's *Voices Under the Window*5 in aid of his analysis of the phenomenology of race and class, all of which make use of hermeneutical and analogical reasoning to help illuminate moral experience via these artworks.

Still other philosophers use "moral psychology" to refer to the scientific study of the determinants of well-being and moral life, often aiming to validate or falsify the presuppositions of reflective and humanistic moral psychology. The spectacular advances of the last fifty years in the cognitive and behavioral sciences fuel this work, which we might call *experimental moral psychology*. There is also ample historical precedent for it in the work of Hobbes, Hume, and Dewey, who all introduce experimental methods of reasoning into moral philosophy. Proponents of experimental moral psychology often can be found enlivening the question-and-answer period of reflective and humanistic conference presentations with pointed questions about the empirical adequacy of various assumptions. (After one recent talk of mine on the topic of emotional self-knowledge, a prominent proponent of this methodology raised his hand to ask, simply, "You know that your conclusion contradicts the science on this matter, right?")

These approaches are, on the whole, largely complementary. Nussbaum's

work on emotion, for example, appeals alternatingly to Aeschylus and to neuroscientist Joseph LeDoux. Likewise, Nietzsche combines literary and historical reflection on Greek tragedy, Christianity, and Wagner with appeals to contemporary (nineteenth-century) psychology, physiology, and race theory (though he rejects certain tendencies of British empirical psychology).

Alfano's introductory text for advanced undergraduate students brings experimental moral psychology into the classroom. It has many virtues. For example, it sketches solutions to numerous controversies in the field, including: (1) whether empirical findings about the indeterminacy and instability of preferences undermine assessments of rightness and well-being, (2) whether findings in neuroscience related to the "dual-process theory" of cognition undermine assessments of rightness that depend on the doctrine of double effect, (3) whether "situationist" findings in social psychology undermine character-based assessments of moral worth and rightness, and (4) whether anthropological findings about the extent of ethical disagreement undermine some versions of moral realism. It also serves as an excellent annotated bibliography of recent experimental moral psychology.

Alfano insists, and I agree, that interdisciplinary confrontation is necessary for doing moral psychology well, since "moral philosophy without psychological content is empty, whereas psychological investigation without philosophical insight is blind" (1). From where should moral philosophers source our psychological content? One answer, which we might call *the broad experimental view*, prioritizes

equally the results of all relevant natural and social sciences, including sociology, criminology, anthropology, archaeology, and linguistics. Even though Alfano does not discuss most of these fields—and it would be practically impossible for any short introductory text to do so—he mentions all of them, at least in passing, in a way that makes clear that he endorses broad experimentalism. Elsewhere Alfano has outlined a research program that is "naturalistic" in the sense that it aims to "employ only methods consonant with those used in the ['hard' and 'soft'] sciences and refer only to entities countenanced by the sciences."6 Alternately, we might accept *a humanistic view*, according to which one necessary component of the psychological content needed for moral philosophy can be acquired only by humanistic study. Such study involves making sense of (for instance) our relation to the past, our aesthetic practices, and our membership in political communities, where the project of sense-making depends crucially—in ways that distinguish it from some scientific projects—on our distinctive cultural, ethical, political, and aesthetic experiences and values.

One point in favor of the humanistic view, it seems to me, is that the study of moral psychology must attend to what our moral experience is like, shaped as it is by our interpretations of that experience. Much moral psychology aims to help us to make sense of our moral experience by using interpretive methods proper to humanistic study, not merely to diagnose the network of causes that make morality "work" (1) or not. Alfano's exclusion of humanistic moral psychology makes me wonder about the place of such reflection in the work of experimental moral psychologists. In their view, must we deny that humanistic reflection

provides necessary psychological content to moral philosophy? Or must humanistic reflection proceed by methods consonant with those of the sciences?

I worry that by ignoring humanistic moral psychology, this introductory text erases much interesting research in moral psychology in that it encourages students to think that there's nothing to see there. Perhaps this is an inadvertent result of, say, marketing pressures from the publisher. Alfano does suggest, though, that the text's design is to offer "a comprehensive survey of contemporary moral psychology" (ix). If so, then Alfano has concealed from the introductory student the fruitful pluralism of the field.

Perhaps Alfano will respond that the book doesn't need to "give the other side," because it *is* the other side. Or perhaps Alfano denies that humanistic moral psychology is intellectually reputable. I look forward to his clarification on this matter.

Implicit Bias

Let's consider one of the many stimulating discussions in Alfano's text—a discussion about blaming people for being implicitly biased that is representative of the mix of empirical and conceptual claims offered in the text. In anti-racist communities there has long been a suspicion, to say the least, that even individuals who sincerely disavow racism might still harbor racial bias. Recent empirical research seems to confirm this suspicion, and there is a burgeoning philosophical literature about how to understand moral responsibility for implicit bias, which Alfano ably

highlights.7 How should we react to people with implicit biases who sincerely avow a commitment to the moral and political equality of all? There are numerous dimensions of analysis that are relevant to answering this question, including (1) whether holding one responsible would be useful (*the pragmatic factors*), (2) whether the biased person is in a position to know about their bias (*the epistemic factors*), and (3) whether the biased person is in a position to control their bias (*the control factors*). Alfano discusses them all, but let us focus on the pragmatic factors.

Alfano presents the following argument that we should *not* think of implicitly biased though explicitly egalitarian persons as racist or sexist. Studies suggest that one's self-conception is often self-confirming: if I conceive of myself as a racist, say, then I am more likely to act like a racist than I would otherwise be. Likewise, accusing others of racism risks making them even more likely to act like a racist. Alfano takes these claims to support what he calls "the factitious, interactionist framework" (132) of virtue, according to which behavior is explained in terms of, among other things, "the ongoing feedback between the individual and environment" (187). On this view, social expectation-signaling and one's self-conception help to bring about and sustain dispositions to think, feel, and act that are similar to traditional Aristotelian virtues or vices. Therefore, conceiving of people as racist is "dangerous" (71). Instead, Alfano suggests, perhaps one should think of such a person "as someone who strives to be fair to targets of negative stereotypes but who suffers in his human, all-too-human, way from various biases" (71). In doing so, one would ascribe lack of ill will (or perhaps even good will) to the

implicitly biased person, which might help to bring about personal improvement. Lest one think that Alfano is urging emotional calm in the face of injustice, he adds that there is reason for victims of implicit bias to "angrily denounce people who are trying their best, despite implicit biases. . . . Even if it ruffles a few feathers" (78). For implicit bias can cause "immense harm" (78).

Several aspects of this argument would make for interesting classroom debate. First, there is some tension between the claims that (a) we should treat such implicitly biased people as if they have genuinely good will (or, at least lack ill will) towards the targets of their bias and (b) we should sometimes angrily denounce such people. For anger and blame are, plausibly, responses to ill will. Pragmatic arguments to be angry in such cases then seem to counsel emotional dishonesty or confusion. Perhaps, as Nussbaum has recently argued, other emotional attitudes are both more fitting and more productive (though perhaps not).[8]

Second, if even angry denunciation is compatible with constructive efforts to improve the offender, then I see no reason why the same cannot be said of calm and supportive communication that the offender embodies some form of racism. Indeed, George Yancy has recently hypothesized that bringing a person's racism to their attention might be crucial for facilitating improvement.[9] Yancy urges well-intentioned whites to consider themselves precisely as people who harbor racism despite their anti-racist intentions, beliefs, and actions. He conceives of such direct communication as a kind of gift, designed to help well-intentioned whites escape from the racist lies and self-deceptions that

cloud their minds. No doubt such a direct intervention requires tact and perhaps even, as Yancy insists, a kind of love. So, on what empirical grounds should we reject Yancy's model of direct anti-racist (and anti-sexist) intervention in favor of Alfano's model?

This question leads to a third concern. It is not clear from the text what empirical support, if any, there is for the claim that the relevant virtue of justice is best characterized as factitious and interactionist. Alfano states that there is *some* evidence that the virtues of tidiness, charity, cooperativeness and competitiveness, helpfulness, eco-friendliness, and scholastic motivation can be understood in this way. In general, though, "we currently lack evidence one way or the other about which virtues" (132) can be thought of in this interactionist way. So, it seems, we don't know whether the virtue of justice at issue is among those that can be inculcated "by fine-tuning your self-concept and the social expectations directed at you" (132). Given this lack evidence, the status of these pragmatic claims about redressing bias is unclear. Does Alfano present them as hypotheses or as justified by particular empirical research? I worry that the reader will receive the false impression that it is a robust finding that we should *not* call (merely) implicitly biased people "racist" or "sexist," and that such a reader will be motivated to adjust their behavior in a problematic direction, for instance away from Yancy's proposal, because of an inadequately supported claim.

Alfano's discussion of the pragmatics of blaming implicitly biased people would make for stimulating classroom debate. That said, this text would work best when supplemented by secondary literature that addresses Alfano's conceptual and empirical assumptions. To this end, Alfano

helpfully provides suggested further readings at the end of each chapter. With respect to the discussion about implicit bias, supplemental readings might engage assumptions about the nature of good and ill will, the ethics of expressing anger and blame, the nature of racism, and the empirical status of Alfano's interactionist theory of virtue.10

4.10.18 | **Mark Alfano**

Reply
Reply to Fischer

Thanks to Jeremy Fischer for his thought-provoking comments on *Moral Psychology: An Introduction.* I welcome this opportunity to reflect on the methodology of moral psychology.

Fischer distinguishes three methodologies in moral psychology, which he dubs *reflective, humanistic,* and *experimental.* Each of these provides its own perspective on the topic. Reflective moral psychology employs armchair pondering of everyday experience, supplemented by common sense. As Fischer notes, reflective moral psychology faces several challenges, and we could easily add more challenges to his list. Which everyday experiences should one reflect on, and why? How universal is common sense? Whose experiences are liable to be ignored in this

0
Shares

f

🐦

♥

process? If someone disagrees with you about a philosophical intuition, does that mean at least one of you lacks common sense?

These are age-old questions, and they have rightly prompted philosophers to seek a wider, more diverse range of experiences on which to reflect and a process that is—in the ideal case, at least—reproducible and intersubjectively valid. This leads us to humanistic and experimental moral psychology. In the former, the range of experience is broadened by going back in time to exemplary historical and literary cases, which furnish rich portraits of people's conduct and inner lives. In this connection, Fischer points to examples such as Bernard Williams's interpretation of *Ajax* and Martha Nussbaum's interpretation of *Oresteia*.

By contrast, experimental moral psychology works with data to seek out trends, identify effects both weak and strong, and construct causal or computational models of moral psychological processes. Instead of uncovering the deep meanings embedded in great historical and mythical exemplars, experimental moral psychology aggregates and analyzes data from ordinary human animals. Fischer questions whether this approach is sufficient. Can an experimental moral psychology reveal everything that a humanistic approach would illuminate? To get a firmer grip on this question, it's helpful to ask what, exactly, distinguishes the humanistic approach. Fischer primarily associates it with *interpretation* or sense-making. Science can tell us what there is and how it works, but humanistic inquiry excels in telling us (or helping us tell ourselves) what it means. I would add that, in engaging our imaginative capacities, humanistic inquiry may also prompt us to consider possibilities and

prospects that have hitherto remained unrealized. Reading Ursula K. Le Guin's stories, for example, is likely to lead to this phenomenon.

Because it can be hard to engage fully with the statistics, humanistic inquiry can be uniquely instructive. Humanistic interpretation explores stylized exemplars, even as it elides the full range of variance in people's psychologies. This has the advantage of focusing our attention on particular cases that make various processes, experiences, and attitudes salient. Just as it is often pedagogically, cognitively, or communicatively helpful to do geometry with drawings on paper, chalkboards, or computer screens, so it is often pedagogically, cognitively, or communicatively helpful to do moral psychology with narratives from books and film. As Andy Clark (2002) has argued, human minds operate best when they are able to iteratively alternate between cognitive processes such as deciding, inferring, and evaluating and perceptual and agential processes such as seeing, feeling, and manipulating. Shifting back and forth between the analysis of data and artistic or hermeneutical representations of idealized types enables us to take advantages of both our cognitive powers and our perceptual and storytelling capacities.

I hope that these irenic remarks go some way to reducing the distance between Fischer and me. However, I don't want to pretend that we agree about everything, for I also see two substantial dangers in the kind of humanistic moral psychology he enjoins. First, the range of exemplars on which to reflect is liable to be at least as cramped as one's own experience. The store of literary and historical exemplars is almost laughably narrow. In works of humanistic moral psychology, how often does one encounter

exemplars who are not "the ancients," characters in Jane Austen or George Eliot novels, or Huckleberry Finn? If one of the deficiencies of reflective moral psychology is that it does not draw on a sufficiently diverse range of actual and possible experience, then humanistic moral psychology is liable to exacerbate the problem. This is not essential to humanistic moral psychology, but one is hard-pressed to find counterexamples. Experimental moral psychology does better on this score (if and only if it uses large and diverse samples of participants—a desideratum that is sometimes neglected).

In addition to (typically) providing a worm's-eye view of a very small number of exemplars, humanistic moral psychology unconstrained by data is prone to lead to *misimagination.* As Adam Morton (2006) points out, it's only possible to learn from our imaginings if they have the possibility of being either correct or incorrect. Fictional, mythical, and idealized historical exemplars are, in this context "an invitation to illusion . . . because when we respond to fiction we react to the characters in many of the ways we do to real people, and so if a way of reacting makes sense with respect to a fiction we tend to think that it makes sense with respect to real people." As Morton goes on to argue, this can lead us to think that the characters, motives, and experiences we attribute to fictional characters are possible (perhaps even desirable) characters, motives, and experiences in real life. It can also lead us to expect causal or conceptual connections in real life that only exist in fiction.

0
Shares

f

𝕏

♥

To illustrate: Bernard Williams moves from shame in Sophocles's *Ajax* to shame in actual human communities. Martha Nussbaum moves from

anger and forgiveness in Aeschylus's *Oresteia* to anger and forgiveness in real human encounters. Gabriele Taylor moves from vice in *Coriolanus* to vice in everyday life. And Charles Mills moves from the experience of class and race in Hearne's *Voices Under the Window* to the experience of class and race in contemporary society. We must ask, though: does shame really work that way? Can real people forgive in the way Nussbaum imagines forgiveness to work in the *Oresteia*? What is vice like for the fancy apes that we are, and how similar is it to the vice of fictional characters in Shakespeare's plays? Do victims of our current racist and classist society experience their own lives as Mills imagines the characters in Hearne's novel to experience them? Maybe. But then again, maybe not. The only way to answer these questions is to employ the methods of science. If my arguments here are on the right track, then humanistic moral psychology may help us make sense of our experience, but it may also help us make nonsense of it. That does not make humanistic moral psychology useless. It does, though, show that the truth-values of the insights it promises need to be corroborated by science. Or, as I put it in the book, "moral philosophy without psychological content is empty, whereas psychological investigation without philosophical insight is blind."

I turn now to Fischer's remarks about bias and responsibility (about which I have more to say in my responses to Trujillo and Radke). Fischer points out that there is a prima facie tension between treating people who embody implicit but not explicit bias as if they have good will, on the one hand, and sometimes angrily denouncing them, on the other hand. I

believe that this tension can be resolved by appealing to the distinction I articulate in more detail in chapter 4 between ascriptions of traits (e.g., "You are such a sexist pig!") and evaluations of actions (e.g., "That was a sexist thing to do!"). The former, but not the latter, tend to function as self-fulfilling prophecies. For this reason, it can be dangerous to accuse other people of harboring biases (even if the accusation is correct), but it is still perfectly possible to get angry and express that anger by calling out bad behavior. Fischer rightly points out that the extant evidence for self-fulfilling prophecies relates to ascriptions of traits other than being a racist or being a sexist. So my tactical advice here and in the book is based on the speculation that the same effect is liable to crop up in this context as well. Further research could corroborate or falsify this speculation. Further research could also shed light on whether my precautionary advice is better tailored to the case than George Yancy's suggestion.

I conclude by noting that all of this may turn out to be moot if the implicit bias paradigm is overthrown. The implicit association test itself may be unreliable and not useful for predicting behavior. I raised this possibility in a tentative way on page 66 of *Moral Psychology*, but recent research has made me even more worried (e.g., Forscher et al. 2017). If implicit bias turns out not to exist or not to have a serious influence on people's conduct, then the conversation we need to have is not about the unicorns who embody implicit-but-not-explicit bias, but about individuals who harbor and express good old-fashioned explicit bias. Indeed, we need to have that conversation no matter what.

References

Alfano, M. 2016. *Moral Psychology: An Introduction.* Malden, MA: Polity.

Clark, A. 2002. "Towards a Science of the Bio-Technological Mind."
International Journal of Cognition and Technology 1: 21–33.

Forscher, P., et al. 2017. "A Meta-Analysis of Change in Implicit Bias."
Open Science Framework, October 5. Retrieved from osf.io/awz2p.

Morton, A. 2006. "Imagination and Mmisimagination." In *The Architecture
of the Imagination: New Essays on Pretence, Possibility, and Fiction,* edited by
S. Nichols. Clarendon.

Add your voice and join the discussion

4.17.18 | **G. M. Trujillo, Jr.**

Response

The Backfire Effect and Political Psychology

Mark Alfano's *Moral Psychology: An Introduction* palatably samples topics in

0
Shares

f

🐦

❤️