T

A Plague on both your Houses: Virtue Theory after Situationism and Repligate

Mark Alfano

Introduction

Over the last two and a half decades, philosophers have failed to come to grips with the implications of social and personality psychology for virtue theory. In this paper, I indict both the critics and the defenders of virtue theory. Critics have relied too heavily on a biased sample of the available evidence. What's more, the most troubling studies and effects that they tend to emphasize have turned out to replicate poorly. At the same time, most virtue theorists have responded not by pointing to these very real flaws in the empirical base for skepticism about character traits, but by retreating into unfalsifiable obscurantism. If my criticisms are on the right track, then virtue theorists have countered bad arguments with worse. I conclude by pointing to more promising directions to follow in theorizing about virtues, vices, and character.

1. *Building on a foundation of sand*

Starting with Owen Flanagan's *Varieties of Moral Personality* (1993), philosophers began to worry that empirical results from social psychology were inconsistent with the structure of human agency presupposed by virtue theory – or at least the neo-Aristotelian virtue theory predominant at the time. In this framework, people are conceived as having more or less fixed traits of character that systematically order their perception, cognition, emotion, reasoning, decision-making, and behavior. For example, a generous person is inclined to notice and seek out opportunities to give supererogatorily to others. The generous person is also inclined to think

about what would (and wouldn't) be appreciated by potential recipients, to feel the urge to give and the glow of satisfaction after giving, to deliberate effectively about when, where, and how to give to whom, to come to firm decisions based on such deliberation, and to follow through on those decisions once they've been made. Other traits are meant to fit the same pattern, structuring perception, cognition, motivation, and action of their bearers. Famous results in social psychology, such as Darley & Batson's (1973) Good Samaritan experiment, seem to tell against this view of human moral conduct. When someone helps another in need, they may do so simply because they are not in a rush, rather than because they are expressing a fixed trait like generosity or compassion.

One might respond by emphasizing that virtue theorists don't have to be optimists; they can explain failures of generosity and compassion by attributing akrasia, enkrasia, or vice rather than virtue (Bates & Kleingeld 2017). In the virtue theoretic framework, people are not necessarily assumed to already be virtuous. However, they are assumed to be at least potentially responsive to the considerations that a virtuous person would ordinarily notice and take into account. Flanagan (1993), followed by Doris (1998, 2002), Harman (1999, 2000), and Alfano (2013), made trouble for this framework by pointing to social psychological evidence suggesting that much of people's thinking, feeling, and acting is instead predicted by (and hence responsive to) situational factors that don't seem to count as reasons *at all* – not even bad reasons or temptations to vice. These include influences such as ambient sensibilia (sounds, smells, light levels, etc.), seemingly trivial and normatively irrelevant inducers of positive and negative moods, order of presentation of stimuli, and a variety of framing and priming effects, many of which are reviewed in Ross & Nisbett (1991) and Alfano (2013, pp. 40-50). It's worth emphasizing the depth of the problem these studies pose. It's not that they suggest that most people aren't virtuous (although they do suggest that as well). It's that they undermine the entire framework in which people are conceived as cognitively sensitive and motivationally responsive to reasons. Someone whose failure to act virtuously because they gave in to temptation can be understood in the virtue theoretic framework. Someone whose failure to act virtuously because they'd just been subliminally primed with physical coldness, which in turn is metaphorically associated with social coldness, finds no place in the virtue theoretic framework. These sorts of effects push us to revamp our whole notion of agency and personhood (Bargh 1999; Doris 2009).

The saving grace of all this, though, is that precisely the most troubling

studies and effects – in which seemingly trivial and normatively irrelevant situational factors predict and explain people's thought, feeling, and behavior better than personality or traits – replicate either poorly or not at all. The replication crisis is an ongoing development within psychology, so it is not yet possible to say definitively which studies do and which do not replicate, but the "Many Labs" collaborations (among other replication efforts) seriously undermine any confidence we might have in the robustness and even existence of a wide variety of framing, priming, and embodied metaphor effects (Klein et al. 2014; Klein et al. 2017; Ebersole et al. 2016). While this does not mean that psychology prompts no revisions to our conception of how virtues are best acquired, maintained, or expressed, it does mean that overall framework is not under threat. Yet, for the most part, defenders of virtue theory have not taken comfort in this fact, choosing instead to shift towards unfalsifiability.

2. *Virtue as a god of the gaps*

In the previous section, I argued that the most troubling evidence for the virtue theoretic perspective on human agency is also among the least replicable. This, I want to suggest, is what virtue theoreticians should have been arguing as psychology's replication crisis began to heat up. While a few philosophers have made efforts to engage the empirical literature as a whole rather than cherry-picking a few studies (e.g., Snow 2009, Russell 2012, Miller 2015), their work has its own problems. For example, Snow (2009) leans heavily on the work of John Bargh, much of which has failed to replicate. Russell (2012) explicitly states that the theory of virtue he endorses is unfalsifiable. Both Snow and Russell also rely heavily on the cognitive-affective personality system (CAPS) model. This is a framework developed in the context of pathological development (at-risk youth), not normal adults; it's hard to imagine that it would ground a theory of *virtue*. Miller (2015) also relies on Bargh's studies, along with a large number of studies from the 1970s and 1980s that have very low statistical power (as few as $n=20$ per cell) and are most likely instances of p-hacking or HARKing (Kerr 1998).

Setting aside these exceptions, most philosophers responding to the empirical challenge to virtue ethics have avoided scrutinizing and engaging with the evidence itself, preferring instead either to argue that it was irrelevant *in principle* or to come up with post hoc stories about particular experiments that had been cited as exemplars of the empirical literature.

Those who argue that empirical evidence is in principle irrelevant end up committed to accepting anecdotal reports of virtue possession and expression while denying that such reports could be systematically collected and analyzed. Those who argue that the (as we now know, flawed) empirical evidence cited by Flanagan, Doris, and Harman happens to be irrelevant typically end up jerry-rigging their accounts of virtue in an effort to make them unfalsifiable. The amount of ink spilled parsing the interpretation of a few seminal studies like the Good Samaritan experiment could fill a lake, even while ongoing and systematic developments (and opportunities for interdisciplinary collaboration) in psychology and interdisciplinary philosophy-cum-psychology have been neglected[1].

In so doing, defenders of virtue ethics have resorted to a strategy analogous to the "God of the gaps" arguments offered by creationist opponents of Darwin. The phrase derives from the self-critical Christian preacher, Henry Drummond, who, in his Lowell Lectures on *The Ascent of Man*, accused his coreligionists thusly:

> There are reverent minds who ceaselessly scan the fields of Nature and the books of Science in search of gaps – gaps which they will fill up with God. As if God lived in gaps! What view of Nature or of Truth is theirs whose interest in Science is not in what it can explain, but in what it cannot, whose quest is ignorance, not knowledge, whose daily dread is that the cloud may lift, and who, as darkness melts from this field or that, begin to tremble for the place of his abode? (Drummond 1894, p. 333).

A century later, one could replace 'God' with 'virtue', 'Science' with '*Science*', and 'Nature' with '*Nature*', and the accusation would ring just as true. Instead of following the best evidence where it leads, philosophers have tended to seek ways to insulate their favored views against any evidence whatsoever. Such a desperate rearguard maneuver is hopeless, and in this case it is also unnecessary. The philosophical response should have been, and still could be, to engage both seriously and critically with the relevant science.

3. *After the plague*

In my more recent work, I've attempted to engage in precisely this way. For example, in Alfano (2016, chapter 4) I argue that the best, aggregated

---

[1] Examples include Annas (2003), Kamtekar (2004), Kristjansson (2008), Sreenivasan (2002, 2008), Upton (2009), and many others.

evidence indicates that eight situational factors explain approximately twice as much of the variance in human behavior as the five main trait factors (Rauthmann et al. 2014). These situational factors (the so-called DIAMONDS) are:

— *duty*: a job must be done;
— *intellect*: the situation affords a chance to demonstrate one's intellect;
— *adversity*: one reacts either prospectively or retrospectively to blame;
— *mating*: one modulates one's behavior because potential romantic partners are present;
— *positivity*: the situation is potentially enjoyable;
— *negativity*: the situation is potentially unenjoyable or anxiety-provoking;
— *deception*: it is possible to deceive someone; and
— *sociality*: social interaction is possible.

Together, these eight kinds of situational influences account for a larger amount of the variance in people's behavior (24-74 percent) than trait dimensions (3-18 percent). Notice, however, that all eight of these dimensions name aspects of situations that provide reasons for thought, feeling, and action. Unlike the faulty evidence on priming, framing, and embodied metaphors, they indicate considerations that, in the virtue-theoretic framework, count in favor of or against having an array of beliefs and motives, in favor of or against undertaking a range of actions and omissions. This sort of evidence may (and probably will) force us to reconsider which temptations we are most prone to, which bad reasons have a tendency to loom too large in our decision-making and policy, which good reasons we have a tendency to neglect. It may help us to formulate a virtue theory that better answers to the types of animals that we are. However, it does not undermine – indeed, it corroborates – the picture of human agency presupposed by virtue theory.

Further research will be needed to map out the details. Perhaps we will end up skeptical of the existence or robustness of certain particularly demanding virtues such as honesty while confident in the existence and robustness of other traits. For example, Fleeson (2001), Fleeson & Gallagher (2009), and Fleeson et al. (2014) provide highly suggestive evidence that most people's patterns of behavior are, though predictable, at best candidates for low- or medium-fidelity traits (virtues, vices, or neither). Some traits predict extremely important and valuable long-term outcomes. For example, people who score low in Propriety – a dimension of the "Big Six" personality model formulated by Saucier (2009) – are much more likely to, at least once in their lives, engage in such morally questionable behaviors

as drunk driving, bar brawls, shoplifting, vehicle theft, assault, and delinquent gang activity (Simms 2007). Along these lines, Jayawickreme et al. (2014) contend that the Agreeableness and Conscientiousness dimensions of the "Big Five" personality model are associated with such low-fidelity virtues as compassion and prudence.

Future research should follow up on these results by developing philosophically sophisticated and empirically validated measures of various candidate virtues (and vices), then using those measures to predict and explain – as best as can be managed – relevant thought, feelings, and behavior. In Alfano et al. (2017) my collaborators and I make a first attempt to do precisely this for the virtue of intellectual humility. Further interdisciplinary research could do likewise for the full range of dispositions in the ethical and epistemic canon.

References

Alfano M. (2013), *Character as Moral Fiction*, Cambridge University Press,Cambridge.

Alfano M. (2016), *Moral Psychology: An Introduction*, Polity, Cambridge.

Alfano M. et al. (2017), *Development and validation of a multi-dimensional measure of intellectual humility*, in «PLoS ONE», 12 (8): e0182950.

Annas J. (2003), *Virtue ethics and social psychology*, in «A Priori», 2, pp. 20-59.

Bargh J. (1999), *The cognitive monster: The case against the controllability of automatic stereotype effects*, in S. Chaiken, Y. Trope (eds.), *Dual-Process Theories in Social Psychology*, Guilford Press, New York, pp. 361-382.

Bates T., Kleingeld P. (2017), *Virtue, vice, and situationism*, in N. Snow (ed.), *Oxford Handbook of Virtue*, Oxford University Press, New York.

Darley J., Batson C.D. (1973), *"From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior*, in «Journal of Personality and Social Psychology», 27, pp. 100-108.

Doris J. (1998), *Persons, situations, and virtue ethics*, in «Nous», 32 (4), pp. 504-540.

Doris J. (2002), *Lack of Character Personality and Moral Behavior*, Cambridge University Press, Cambridge.

Doris J. (2009), *Skepticism about persons*, in «Philosophical Issues», 19 (1), pp. 57-91.

Drummond H. (1894), *The Ascent of Man*, James Pott & Co. Publishers, New York.

Ebersole C. et al. (2016), *Many Labs 3: Evaluating participant pool quality across the academic semester via replication*, in «Journal of Experimental Social Psychology», 67, pp. 68-82.

Flanagan O. (1993), *Varieties of Moral Personality: Ethics and Psychological Realism*, Harvard University Press, Cambridge.

Fleeson W. (2001), *Toward a structure- and process-integrated view of personality: Traits as density distributions of states*, in «Journal of Personality and Social Psychology», 80, pp. 1011-1027.

Fleeson W. et al. (2014), *Character: The prospects for a personality-based perspective on morality*, in «Social and Personality Psychology Compass», 8 (4), pp. 178-191.

Fleeson W., Gallagher P. (2009), *The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis*, in «Journal of Personality and Social Psychology», 97, pp. 1097-1114.

Harman G. (1999), *Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error*, in «Proceedings of the Aristotelian Society», New Series, 119, pp. 316-326.

Harman G. (2000), *The nonexistence of character traits*, in «Proceedings of the Aristotelian Society», 100, pp. 223-226.

Jayawickreme E. et al. (2014), *Virtuous states and virtuous traits: How the empirical evidence regarding the existence of broad traits saves virtue ethics from the situationist critique*, in «Theory and Research in Education», pp. 1-26.

Kamtekar R. (2004), *Situationism and virtue ethics on the content of our character*, in «Ethics», 114, pp. 458-491.

Kerr N. (1998), *HARKing: Hypothesizing after the results are known*, in «Personality and Social Psychology Review», 2 (3), pp. 196-217.

Klein R. et al. (2014), *Investigating variation in replicability: A "Many Labs" replication project*, in «Social Psychology», 45 (3), pp. 142-152.

Klein R.A. et al. (2017), *Many Labs 2: Investigating variation in replicability across sample and setting*, in *Advances in Methods and Practices in Psychology Science* (RRR).

Kristjansson K. (2008), *An Aristotelian critique of situationism*, in «Philosophy», 83, pp. 55-76.

Miller C. (2015), *Moral Character: An Empirical Theory*, Oxford University Press, Oxford.

Rauthmann J. et al. (2014), *The situational eight DIAMONDS: A taxonomy of major dimensions of situational characteristics*, in «Journal of Personality and Social Psychology», 107 (4), pp. 677-718.

Ross L., Nisbett R. (1991), *The Person and the Situation*, Cambridge University Press, Cambridge.

Russell D. (2012), *Practical Intelligence and the Virtues*, Oxford University Press, Oxford.

Saucier G. (2009), *Recurrent personality dimensions in inclusive lexical studies: Indications for a Big Six structure*, in «Journal of Personality», 77 (5), pp. 1577-1614.

Simms L. (2007), *The Big Seven model of personality and its relevance to personality pathology*, in «Journal of Personality», 57, pp. 65-94.

Snow N. (2009), *Virtue as Social Intelligence: An Empirically Grounded Theory*, Routledge, New York.

Sreenivasan G. (2002), *Errors about errors: Virtue theory and trait attribution*, in «Mind», 111, pp. 47-68.

Sreenivas G. (2008), *Situations against virtues: The situationist attack on virtue theory*, in Mantzavinos (ed.), *The Philosophy of Social Sciences: Philosophical Theory and Scientific Practice*, Cambridge University Press, New York, pp. 274-291.

Upton C. (2009), *The structure of character*, in «The Journal of Ethics», 13, pp. 175-193.

Abstract

Virtues are dispositions that make their bearers admirable. Dispositions can be studied scientifically by systematically varying whether their alleged bearers are in (or take themselves to be in) the dispositions' eliciting conditions. In recent decades, empirically-minded philosophers looked to social and personality psychology to study the extent to which ordinary humans embody dispositions traditionally considered admirable in the Aristotelian tradition. This led some to conclude that virtues are not attainable ideals, and that we should focus our ethical reflection and efforts more on jerry-rigging our environments than on improving our characters. Most virtue ethicists resisted this reorientation. However, much of the scientific evidence on which the controversy was based has failed to replicate, raising the question of how much faith we should place in methodologically suspect studies. In this paper, I assess the state of the debate and recommend best practices for a renewed interdisciplinary investigation of virtues and vices in which philosophical expertise related to conceptualization and theorizing is essentially intertwined with scientific expertise related to operationalization, measurement, and statistics.

Mark Alfano
Associate Professor, Delft University of Philosophy
Professor, Australian Catholic University
mark.alfano@gmail.com