Forthcoming in *The Oxford Handbook of Moral Psychology*

Situationism, Moral Improvement, and Moral Responsibility

MARIA WAGGONER, JOHN M. DORIS, AND MANUEL VARGAS

Abstract: In this chapter, we recount some of the most pressing objections to character scepticism, pointing out their limitations and, when appropriate, incorporating their suggestions. From here, we consider what empirically informed moral improvement might look like by turning to the skill analogy. While the skill analogy provides a realistic rubric for becoming a better person, many of the questions concerning the details of how moral improvement might take place remain unanswered. When developing expertise in domains like chess and morality, a wide range of factors will likely be important and it is unlikely than any one individual factor will be especially important. Given this, for any account of moral improvement, our optimism should be bounded: the effect of any particular intervention is likely to be limited, in both magnitude and domain. Lastly, we consider how character scepticism has reshaped the way we think about moral responsibility, whether this is cashed out in reasons-responsiveness or self-expression accounts. While both views face challenges, by distinguishing the possession of the responsibility-making feature from failures to manifest that property in behaviour, we gain a certain degree of wiggle room that allows us to accommodate empirical findings while holding on to notions of responsibility. This chapter highlights that the 'person-situationism debate' has expanded far beyond its beginnings, giving way to new accounts of moral improvement and moral responsibility.

Keywords: moral psychology, character, situationism, virtue ethics, moral improvement, skill analogy, moral responsibility, moral agency

32.1 Introduction

Starting in the 1990s, philosophers inspired by 'person-situationism debate' that had unsettled personality and social psychology since the 1960s instigated the 'virtue ethics-situationism debate' concerning the appropriate role for character in philosophical ethics and moral psychology (Alfano 2013; Doris 1998; 2002; 2005; 2010; forthcoming; Harman 1999; 2000; 2001; 2003; 2009; Machery 2010; Merritt 2000; Merritt, Doris, and

Harman 2010; Vranas 2005).[1] Much as 'situationist' social psychologists evinced scepticism about the importance of personality traits in the explanation and prediction of behaviour, philosophical 'character sceptics' contended that the characterological moral psychology typical of neo-Aristotelian virtue ethics—resurgent in moral philosophy since the 1950s (e.g. Anscombe 1958; Foot 1978)––is 'empirically inadequate' and fails standards of 'psychological realism' (Flanagan 1991) in ethical theorizing.

The instigating science was the oft-repeated finding that seemingly arbitrary and insubstantial situational factors have rather substantial effects on our behaviour, suggesting that behavioural *consistency* is lower than would be expected if behaviour is typically ordered by 'global' or 'robust' character traits like virtue and vices. The scientific literature is by now familiar enough to students of moral psychology to make detailed discussion superfluous for this review chapter, but among the many representative findings are that finding dimes (Isen and Levin 1972: 387) and smelling cinnamon rolls (Baron 1997: 500–501) prompt us to help others, while hot weather (Kenrick and MacFarlane 1986: 184–7; Anderson 2001: 34–6), the noise of a lawnmower (Mathews and Cannon 1975: 574–5), and being in a hurry (Darley and Batson 1973: 105) impede prosocial behaviour.[2] Most famous, and most heavily relied on by character sceptics, are Milgram's (1974) studies of obedience, where people were willing to harm a protesting victim when asked to do so by a guy in a white lab coat.

In light of this evidence, character sceptics have drawn two implications. First, they make a *descriptive* claim: the limited influence of personality variables and the rather surprisingly potent influence of situational variables together suggest that character has a less prominent part in structuring behaviour than character theorists, and 'common sense', suppose. Second—the delicate relationship between the empirical and normative duly noted (Doris and Stich 2005; Railton 2004)––the sceptics usually follow their descriptive claims with *prescriptive* claims about how ethical thought might best proceed. For example, some character sceptics argue that we would do better if we spent less time and

[1] Alston (1975) and Flanagan (1991) are philosophers who early on discussed the relevant literature in psychology, but neither advocated scepticism about character.

[2] For detailed surveys, see Doris (2002); Miller (2013; 2014); Ross and Nisbett (1991); and Vranas (2005).

effort trying to cultivate virtue—an endeavor psychological science suggests is likely to be daunting—and instead focused more on fostering situations, relationships, and institutions conducive to morally optimal behaviour (Doris 2002; Harman 1999; Merritt 2000).

Over the past decade or so, philosophical conversations about the import of situationist social psychology have grown beyond initial treatments of the descriptive and prescriptive issues; particularly lively are discussions about the possibility of moral improvement, and the empirical findings' import for moral responsibility. In what follows, we (1) canvass initial debates about character scepticism, and then consider later developments concerning (2) the implications of character scepticism for moral improvement, and (3) how moral responsibility theory has grappled with situationist findings in social psychology.

32.2 Responses to character scepticism

In this section we canvass the main issues in the original debates about character scepticism. We first consider arguments against the character sceptic's descriptive claim, and then consider arguments aimed at the prescriptive claim.

Some have sought to reject the sceptics' descriptive claim, usually by either discrediting the empirical evidence or arguing that the evidence does not support sceptical conclusions. Here, we consider three versions of these anti-sceptical manoeuvres.

(1) The experimental scenarios are ethically inconsequential, and so don't address moral character (Sabini and Silver 2005: 540; Sreenivasan 2002: 59).

(2) The empirical work in question typically involve 'one-off' rather than intrasubject, longitudinal studies, and so fail to provide information about any behavioural consistency across diverse situations (Fleeson and Furr 2016: 236–8; Slingerland 2011: 395–6; Sreenivasan 2008: 607).

(3) Many of the studies in question, such as the 'dime in the phonebooth' study, are subject to replication concerns, and therefore are devoid of evidential value (Alfano 2018: 115; Miller 2003: appendix; Webber 2006: 653).

With regard to (1), it's arguable that some of the experimental behaviours, like helping someone pick up spilt papers, may be morally unimportant (Alfano 2013: 71). Conversely, it's arguable that such 'small-scale'

behaviours are morally telling (Doris 2005: 662); a callous failure to help remains a moral failure, even if the stakes are not life and death. But even if *some* of the evidence can be dismissed in this way, certainly not *all* of it can be: it's biting a large bullet to call administering seemingly fatal shocks to an innocent person (Milgram 1974), or neglecting a stranger who appeared to be in considerable distress (Darley and Latané 1968; Darley and Batson 1973), morally unimportant—to mention just two of the awkward examples in the experimental literature. Additionally, surprising, often horrific, moral lapses by seemingly decent people are easily found in the historical literature.

More serious, perhaps, is the lack of longitudinal studies in the sceptics' database; a single experimental observation does not speak directly to behavioural inconsistency. (Likewise, a single observation does not speak directly to behavioural *consistency*.) The sceptic is therefore required to make a sort of *indirect* argument. Where an experiment induces substantially counter-normative behaviour—like administering shocks to a screaming victim—the sceptic *infers* inconsistency from the fact that most people do not typically do such things. The comparative ease with which counter-normative behaviours are induced suggest that they are common in naturalistic contexts, especially since 'real-world' situational pressures to counter-normative behaviour may be more substantial than many experimental ones; for example, totalitarian state apparatuses have lamentable success in inducing Milgram-like destructive obedience from their subjects, who may otherwise seem ordinarily upstanding.

The final objection to the empirical evidence for scepticism adverts to the 'replication crisis' that roiled psychology starting around the 2010s (Chambers 2017; Doris 2015: 44–9); the suggestion is that key experiments in the situationist tradition may not be reproducible. The concerns about replication need to be taken seriously, and there's no doubt that some celebrated findings should be celebrated no more. But we should hesitate to conclude that the experiments motivating situationism should be dismissed en masse. For instance, the Milgram studies, arguably the central exhibit in the sceptics' case (Webber 2006: 656), have certainly been replicated, and bystander group effects—another central strand of evidence for scepticism—remain in good standing (see Fischer et al. 2011; Latané and Nida 1981).

There is also a more general reason why replication problems will not undermine character scepticism. It is widely agreed in psychology, by personologists and situationists alike, that effect sizes reflecting the influence of personality on particular behaviours of interest can typically be expected to reach not much more than a correlation of 0.3, with many published findings being considerably smaller (Mischel 1968: 77–8; Roberts Kuncel, Shiner, Caspi, and Goldberg 2007; Ross and Nisbett 1991: 90–118; Sabini and Silver 2005: 540–42). The interpretive matters here are difficult, but the basic point is that correlations of less than around .15 are not usually detectable by 'casual observation'—remember that a correlation of .00 indicates that two variables are unrelated—while a relationship of .3 might best be characterized as noticeable, but not dramatic (for fuller discussion, see Doris forthcoming). That is, on any given occasion, character traits may be expected to have an influence that is at most noticeable, and far from decisive—a rather far cry from 'Character is destiny.'

One important reason for this circumstance is that behavioural outcomes are typically the subject of multiple variables, and where this is the case, the influence of no one variable will be especially large, with a moderate effect size of about .5 being a plausible limit (Ahadi and Deiner 1989: 403). This circumstance, it should be noted, is not limited to personality variables: .3 is a plausible 'soft' upper limit for effect sizes in social psychology, and other areas of psychology as well (cf. Funder and Ozer 2019). While exceptions appear in the literature, the finding that the influence of personality variables is expected to range over small to moderate effect sizes is 'replicated' countless times in labs around the world. Therefore, a central empirical claim for character scepticism—that the influence of character on behaviour is limited—is not subject to replication concerns, even if some of studies that initially motivated character scepticism fail to replicate.

Instead of calling into question the empirical findings themselves, another way to resist character scepticism is to deny that virtue ethics and characterological moral psychology are committed to the kind of empirical claims the evidence problematizes. Even if it's right to think that we exhibit *overt* behavioural inconsistencies, many (Kamtekar 2004; Upton 2009) have argued that this only problematizes 'behaviourist' accounts of character of a sort no virtue theorist actually holds. The virtue ethics tradition emphasizes the agent's inner states, like her emotional

proclivities and rational abilities (e.g. Adams 2006; Swanton 2003), and these (so the objection goes) are not addressed by the situationist psychological studies.

We should notice that this response, insofar as it is offered as a response to situationism, is committed to an empirical claim to the effect that the relevant psychological processes exhibit considerable cross-situational consistency—presumably, more so than does overt behaviour. Yet there is a very substantial empirical literature indicating that many psychological processes are themselves subject to arbitrary situational variation (Doris 2005; 2015; Olin and Doris 2014): here one might advert to the extensive 'heuristics and biases' tradition demonstrating shortcomings of human rationality (Baron 1994; 2001; Gilovich, Griffin, and Kahneman 2002; Kahneman, Slovic, and Tversky 1982; Kahneman and Tverskey 1982; Kruger and Dunning 1999; Nisbett and Borgida 1975; Nisbett and Ross 1980; Stich 1990; Tversky and Kahneman 1981), or the literature on the difficulty people experience 'transferring' problem-solving skills from one domain to another (Ceci 1993a; 1993b). In short, it's not just the consistency of behaviour that the empirical studies call into question, but the consistency of psychological states as well. Much turns, however, on how consistency is to be understood. For instance, Upton (2009: 178) contends that in order to appropriately draw conclusions about the agent's character from exhibited behaviour, the situations in question must be 'individuated from the agent's point of view, rather than from an outsider's'. Yet, very often, social psychologists only have access to a *nominal*, or third-person, perspective on participants, meaning that behavioural measures in their studies will frequently omit how the subject is construing her situation. Therefore, the anti-sceptic contends, many findings from social psychology fail to address the sort of consistency at issue for the character theorist.

While it is true that many of the situationist experimental paradigms do not assess subjects' subjective construals, it's unclear to what extent doing so would impact the character sceptics' conclusions. For one, many people exhibit inconsistency *by their own lights*: a natural reading of the distress and conflict exhibited by Milgram's subjects—and many perpetrators of real-world destructive obedience—is that they were not consistently adhering to their *own* ethical standards (Papish 2017: 542–4).

Moreover, there remains the question of whether 'by their own lights' nominal consistency is the kind of consistency we ought to be

primarily concerned with. Here, descriptive issues intermingle with evaluative ones. Often, we hold people accountable according to universal moral standards or, less ambitiously, by the shared standards of some cultural group, and inconsistently adhering to these shared or impersonal standards is not excusable by noting that the agent is consistently adhering to her own ethical standards (Alfano 2013: 78–9).

Finally, if one's subjective construals are an important aspect of character, then it must be true that being virtuous requires attending to certain features of our environment, or to interpreting our circumstances, in certain ways and not others. Failing to understand the administering of potentially lethal shocks to an innocent person as anything other than something morally abhorrent is itself evidence of a moral failing, a failing that is not explained away by noting that the agent did not take herself to be doing something wrong. Far from explaining away troubling ethical inconsistency, some construals themselves may be ethically culpable.

The upshot, we think all parties will agree, is that *both* inner states and outer behaviour matter. The challenge for moral psychology is to develop theories of character through empirically credible and theoretically useful accounts of how the inner and outer together work to shape human lives.

Another batch of responses focuses not on *dismissing* the empirical evidence, but rethinking our understanding of character traits *in light of* the evidence. One attempt to do so is the 'local trait' theory proposed by various philosophers (Adams 2006; Doris 2002; Upton 2009; Vranas 2005). While behaviour is cross-situationally quite variable, it is often temporally stable over iterated trials of similar situations, and some theorists have attempted to develop this observation into an account of character traits. On this view, while global highly general traits issuing in cross-situationally consistent behaviour are unlikely to be widely instantiated, fine grained, situation-specific dispositions—e.g. beneficence-to-a-close-friend-when-smelling-perfume—might be (Doris 2002: 62–8). These local traits look to be a departure from traditional character theory, since typical trait attributions seem not to carry such fine-grained qualifications: beneficence-to-a-close-friend-when-smelling-perfume doesn't seem to be the stuff of which bards sing. Yet numerous virtue theorists (Upton 2009; Adams 2006; Grover 2012) develop local trait constructs into theories of virtue; local traits, they think, can found a distinctively virtue-theoretic approach to moral psychology and normative ethics, and any loss of

theoretical economy and normative appeal is counterbalanced by gains in empirical adequacy. Rock-climbing-in-reasonable-weather-courageous, for example, while a downsizing of courage simpliciter, is certainly an apt basis for assessment and aspiration, and so may guide our normative thought. The descriptive moral psychology suggested by this approach will of course be less economical than theories featuring more global traits; but in an uncooperative world, simple theories risk empirical inadequacy.

Another way the situationist data might be accommodated is through Merritt's account (2000) of 'socially sustained' virtue, where virtue-appropriate behaviour may only be reliably realized in properly constituted social environments. Similarly, Pettit (2015: 71) puts forth an 'ecological' account of virtue, whereby virtue only develops in a 'suitable social environment'.

These accounts are attractive because they seem amenable to the lessons from social psychology—and, indeed, from all the agonies of human history. Yet, there are questions about the extent to which they depart from, or even overturn, traditional (especially Aristotelean) virtue theory.[3] Annas (2003: 25), for example, is one Aristotelean traditionalist who doesn't welcome Merritt's proposal, as it abdicates the 'robustness' of virtue that lends the tradition a large measure of its appeal. It's attractive to think of the virtuous as those that don't just do good when their environment makes it easy, but also are able to do good *despite* not having a facilitating social infrastructure. This is part of the appeal in thinking that moral dissidents like King and Gandhi are virtuous—they were at their best when virtue was *not* socially sustained. Thus, socially sustained accounts may be seen as departures from tradition. Indeed, rather than counting as a critique of character scepticism, accounts of socially sustained virtue are ones that character sceptics may happily take on board, since they front-load the importance of the kind of situational influence that motivated character scepticism in the first place.

Other alternative theories of character have been developed in order to account for the evidence of our behavioural inconsistencies

[3] Although we here focus on the broadly Aristotelian approaches that have been the focal target of character scepticism, numerous scholars have developed responses to scepticism sourced in Confucian virtue ethics, which emphasizes the importance of social supports for virtue—e.g. the use of rituals to help construct people's circumstances in morally beneficial ways (Mower 2013; Slingerland 2011; Hutton 2006).

without doing away with global traits by appealing to the important explanatory role that (clusters of) mental states play. Snow (2009) and Russell (2009) have employed Mischel and colleague's Cognitive-Affective Personality System model (CAPS): CAPS proposes that human beings have mental networks of situational-input behavioural-output links, such that situational inputs are mediated by, or filtered through, people's idiosyncratic cognitive and affective dispositions. Two people might encounter the same objective, or nominal, situation but, because of their differing cognitive-affective systems, respond in very different ways. Inasmuch as these systems issue in orderly patterns of behaviour, they may be thought of as the underpinnings of global character traits.

However, there is some question about whether the CAPS model can be used as a framework for *moral* character traits. As Papish (2017: 543, n. 13) observes, on CAPS,

> a person's [moral] values are merely one element among the many that mediate between a person and the environment. There is simply no [. . .] reason to conclude that anything resembling a considered moral judgment will be more determinative of how a person responds to a situation than, say, her stereotypical beliefs, her affective responses, or the constructs that ground her self-image.

Yet, if CAPS is to be a model for *moral* character, moral values must have some sort of priority. For instance, in her treatment of CAPS, Snow (2009: 36) suggests that experiencing a conflict between one's moral values and emotions or behaviours will prompt reflection and efforts to change them so as to align with one's values. However, it is far from clear that this is what CAPS would predict, as opposed to, say, continuing to experience conflict, or changing moral values to align with emotions or behaviours. More generally, as Miller (2014: 218; 2017: 467) points out, CAPS is best understood as account of personality organization compatible with various accounts of character traits, rather than itself being an account of character traits. Therefore, while CAPS may be an element in a theory of virtue, crafting it into a full-blown virtue theory would require considerable filling out.[4]

[4] Russell (2009: 323) is clear on this; he is content to argue that a CAPS-based virtue theory is a 'real possibility'.

Miller (2013; 2014) was an early philosophical proponent of CAPS, but has since abandoned it in favour of a theory maintaining that while traditional virtues and vices are largely absent from the human population, most people possess 'mixed traits'. For Miller (2014: 207–9), mixed traits cannot qualify as virtues or vices, because they are not behaviourally uniform; they sometimes issue in behaviour conforming to the conduct featured in their names—mixed aggression, mixed helping, etc.––but at other times they don't. According to Miller, psychological situations facilitate the expression of mixed traits; activation or inhibition of mixed traits by psychological features of the situation effects orderly patterns of morally relevant behaviour, such as those involving helping and failing to help as the actor perceives to be appropriate or expedient. On Miller's understanding of mixed traits, people will consistently behave poorly in some nominal situations while consistently behaving well in others; the result is that most of us are far from virtuous (and far from vicious). Then while Miller's theory is anti-sceptical regarding traits, it can be thought of as somewhat sceptical regarding virtue.

As a descriptive theory in moral psychology, there is some question as to the theoretical wieldiness of Miller's proposal, as there is for local traits, insofar as mixed traits may be highly complex entities involving multiple determinants of behavioural variation. On the normative side, one alleged disadvantage of using mixed traits (e.g. making traits broad but evaluatively inconsistent) rather than local traits (e.g. making traits narrow but evaluatively consistent) is that local traits, but not mixed traits, seem to better approximate everyday practices of moral appraisal (see Slingerland 2011: 402–3; Upton 2009: 186–9).[5] It is not entirely clear what moral psychology 'everyday practice' supposes, nor is it clear that that a philosophical moral psychology must be beholden to it, but it does seem fair to say that mixed traits are not an intuitive foundation for moral assessment.

Finally, Miller's theory raises questions concerning the cultivation of virtue. As Miller makes explicit, mixed traits are not virtues, nor are they materials out of which virtues can be readily constructed, because virtues are expected to be evaluatively uniform, and mixed traits, by definition, are

[5] Note that Aristotle recognized characterological categories typified by inconsistent behaviour—Swanton (2003: 30) and Miller (2003: 379) both remark that his incontinent person is one who may do what is right when it is easy for her, but act in morally inappropriate ways when the going gets tough.

not. If mixed traits theory gives us the right account of personality organization, how can personalities so organized come to realize, or at least better approximate, virtue? Miller (2014: 227–39) terms this question the 'realism challenge', deeming it the most serious difficulty facing virtue ethics, and he offers some preliminary ideas about how to address it. Later, we shall likewise discuss the issue of moral improvement. But first, we turn to some responses to character scepticism's prescriptive program.

Even if character sceptics are correct in their descriptive claims, there is a further question about how we ought to proceed in normative ethics; virtue ethics might turn out to be the most appealing option on offer, even if it must be divested of problematic elements in its associated moral psychology. And, whatever one thinks of virtue ethics, it is arguable that the normative implications of character scepticism are untenable, however perspicuous its descriptive moral psychology.

Unsurprisingly, given their views about the limited behavioural potency of personality traits, some character sceptics (Doris 2002: 147–8; Harman 2003: 91) prescribe that we focus on the situations we place ourselves in. (As we'll see later, this move anticipates an emphasis on shaping environments that also occurs in the literature on situationist responses to responsibility.) Here, securing morally appropriate behaviour becomes less a matter of self-cultivation than situational management: if you think eating meat is immoral, you're better off throwing out the bacon than exercising your will every time you pass the fridge. Yet some have questioned whether such a prescription can be issued by someone who doubts that character exerts decisive influence on conduct: Sarkissian (2010) argues that how we shape our future situations is, itself, a function of our character. Both Rogers and Warmke (2015) and Kleingeld (2015) continue this line of reasoning, arguing that if it is true that situational factors have a rather significant effect on our behaviour, then the character sceptic puts forth unrealistic prescriptions, for any attempt to choose or construct beneficial situations for ourselves will *itself* be subject to situational perturbations. To simplify a bit, tossing out the bacon is as much an exercise of character, or pretty nearly so, as declining to eat it, so the situational management proposed by the sceptic, far from eschewing reliance in character, positively requires it.

The first thing to say is that not all situations are equally challenging: while a systematic theory of 'situational difficulty' is not be in the offing, surely the wavering vegetarian has a better chance of holding

the line at a vegan cafe than a barbecue. Perhaps, however, both sides can be right. It certainly appears as though people can successfully meet normative demands and aspirations; lots of people consistently follow the practice of moral vegetarianism, and other normatively demanding ways of life, despite the omnipresence of situational impediments like bacon. At the same time, it also appears that situational management can carry us through where relying on our character would leave us falling short; 'Stay out of bars' seems good advice for the alcoholic new to recovery. As is universally agreed, conduct is inter alia a function of a 'person x situation *interaction*' (Mehl, Bollich, Doris, and Vazire 2015: 630), and that observation is certainly in force when thinking about securing moral behaviour: however fragile our dispositions are, they can help enable us to 'bootstrap' ourselves into situations which are conducive to their expression. Where we're disposed to act morally, this dynamic—and partly person-directed—interaction can produce morally appropriate behaviour.

Aside from situational management, other effective options may exist—for the smoker who is trying to quit, she might more directly intervene on her desires—which is, arguably, distinct from both situational management and cultivating virtue—by using a nicotine patch. Alternatives such as these have largely been underexplored within the current literature on moral improvement and virtue cultivation, but one proposal that warrants further consideration is Upton's (2017) appeal to the benefits of meditation; there may be reason to see the psychology of one who routinely practises certain kinds of meditation as one who becomes 'immune' to many kinds of situational influence.

Nonetheless, even if we can at least sometimes navigate our current spaces to select or construct better situations for ourselves, it is far from clear if this prescriptive claim is the *only* consequence of character scepticism. While Doris allows that situationism is 'conservatively revisionary', he denies that it is 'radically revisionary', since doing away with character does not entail 'erod[ing] materials required for a viable (and recognizably ethical) ethical practice' (Doris 2002: 129). Against this, D'Cruz and Cohen (2016) contend that casting doubt on the consistency of our behaviour doesn't just give rise to character scepticism, but undermines crucial trust-based interactional practices like promising. If we are convinced by the character sceptic, they contend,

> when a person caves in to situational pressure and fails to act as she promised, we may be disappointed but we will

> not be indignant. We will see her failure [. . .] as the predictable behaviour of a being for whom the normative expectation of cross-situational consistency makes little sense. (D'Cruz and Cohen 2016: 226)

If this argument goes through, character scepticism cuts beyond character, more pervasively undercutting our moral expectations of others.

Plausibly, character scepticism is at least somewhat revisionist. Perhaps, the more scientifically informed a theory of moral personality is, the more likely it is to be revisionary—and the more likely it is to sacrifice normative appeal and practical adequacy. Something of this sort goes on in other areas of science, such as biology or physics: while such disciplines might have started off using ordinary folk terms (e.g. 'life', 'movement', 'space'), with scientific advances, these notions were revised, taking on new meanings which diverge from their folk understanding (Hochstein 2017: 1131–2; Vargas 2013a: 75–7). Nonetheless, character scepticism probably has room for—and may in fact complement—normative practices like trusting others, making promises, and holding others to certain normative expectations. For instance, we might think of promise-making as a pre-commitment device whereby we change the features of our social situations, making them more conducive for carrying out the behaviours detailed in our promises: Kanngiesser, Sunderarajan, and Woike (2020: 1) found that 'promises [to not cheat] systemically lowered cheating behavior'.[6] More generally, some research suggests that we are more likely to achieve our goals when we impose a high cost on ourselves for deviating from them. Ariely and Wertenbroch, (2002) found that students who set costly self-imposed deadlines procrastinated less and performed better on their writing assignments, while Cawley and Price (2011) found that those who wagered their own money on future weight loss were more likely to shed the pounds. It's plausible that making a promise works similarly, by self-imposing costs if we fail to follow through (Charness and Dufwenberg (2006) suggest that guilt aversion motivates promise-keeping). Even if we are overly optimistic about the extent to which we take ourselves to be capable of making good on our promises, the practice of promising certainly seems to work, and arguably better than not having such a

[6] Such effects haven't always been found: pledges to sexual abstinence among adolescents showed no reduction in the number of future sexual encounters (Rosenbaum 2006).

practice, if our concern is that our commitments be kept. Doubtless, it doesn't work as often as we'd like, but promising works, and works well enough to support a robust practice; the lesson according to the situationist, of course, is that this efficacy is substantially due to the support of an external social 'scaffolding'.

Alfano's (2013) account of factitious virtue carries a similar lesson: despite being a *character* sceptic, he isn't quite a *virtue* sceptic (2013: 13), for he thinks the discourse of virtue has an important practical role in securing ethically appropriate behaviour. Labelling others as virtuous can be a way to change their situation, prompting behaviours that appear to be in line with virtue. The general gist is that making moral commitments—whether taking them on ourselves, as in the case of making promises, or placing them on others, as in the case of virtue labelling—can have effects on our resultant behaviour, since the promise or label motivates us to act in ways to uphold it.

Thus far, we've covered varying objections launched against the character sceptic's original descriptive and prescriptive claims. Debate has since evolved into two larger bodies of philosophical discussion: (1) evaluation of the character sceptic's prescription to focus on situational management gave way to broader questions concerning empirically informed approaches to moral improvement, while (2) the descriptive claims that arose from situationist social psychology led many to reconsider what these empirical findings mean for moral agency and responsibility. We examine these issues in the next two sections.

32.3 Moral improvement

Character scepticism alleges that the robust dispositions associated with virtue are seldom instantiated in actual human psychologies––a point numerous defenders of virtue ethics readily acknowledge (Kamtekar 2004: 466; Solomon 2003: 48, 56; Wielenberg 2006: 471–90): from antiquity to the present, many philosophers have contended that virtue is rare. However, even if many people aren't virtuous *yet*, this circumstance does not preclude the possibility of moral improvement, and progress in attaining or approximating virtue.

For example, some philosophers (Adams 2006; Snow 2009) have suggested we can acknowledge that people typically have only, as Upton (2009: 186–9) has suggested, local virtues. But with effort and practice, the argument continues, one can 'expand' local virtues into something

broader, e.g. into global virtues—even if one starts with only 'rock-climbing-courage', one can, with the right sort of practice, expand one's virtue from the crags to the world, and eventually attain unqualified, global courage.

In her account of virtue development, Snow (2009) suggests that while 'our virtues might start out by being local, they need not remain so' (p. 27). In particular, Snow proposes that we can work to change our psychological situations by adopting different construals or appraisals, (pp. 33–4), thereby changing our reactions. Yet it is not clear why any given successful change to one of our construals should be expected to extend to others. In CAPS terms, we shouldn't expect even successful efforts to impact a given construal beyond one, particular cognitive-affective link (e.g. viewing with a more compassionate lens that particular demanding student who comes in during office hours). Thus, successful changes may be highly context-sensitive. Additionally, where construals and appraisals are emotionally laden (Roberts 2013), they may be especially change-resistant (Kurth 2021). Finally, if the suggestion is to be more than an aspiration, we should like detailed instructions as to what makes this exercise effective. And of course, the same demand obtains for any other program of moral improvement—how exactly does the program work?

Recently, the *skill analogy* has been proposed as a rubric for addressing this challenge: with an appropriate regime for developing moral skill, virtue should become more common, and character scepticism itself would thereby be empirically undermined (Lott 2014, Magundayao 2013). Philosophers have long compared moral goodness to an acquired skill,[7] and many contemporary ethicists have endorsed this approach:[8] for instance, Annas (2011: 1) asserts that '[t]he acquisition and exercise of virtue can be seen to be in many ways like the acquisition and exercise of more mundane activities, such as farming, building or playing the piano', while Russell (2015b: 103) declares that the 'cognitive and affect barriers to acquiring virtue are no different from the barriers to learning a complex skill'.

[7] Aristotle did not unreservedly endorse treating virtues as skills (e.g. 1984: 1105a26–b4), perhaps making the Stoics a more likely an inspiration for the skill analogy (Bloomfield 2001: ch.2; but see Stichter, 2007).

[8] E.g. Bloomfield (2000; 2001; 2014); Ciurria (2014); Fridland (2017); Jacobson (2005); Russell (2015a; 2015b); Snow (2009: e.g. 74); Sosa (2009); Stichter (2007; 2011; 2018).

Insofar as doing good can be a challenging exercise, thinking of moral excellence as a sort of expertise is intuitively appealing. Moreover, inasmuch as skill acquisition in non-moral domains is comparatively well studied, the skill analogy may be able to exploit what is known in these domains in an account of moral development. Chess (with the possible exception of musical ability) is perhaps the best-studied skill (Simon and Chase 1973: 394), and chess is often appealed to, as we will do here, by philosophers exploring the skill analogy (Bloomfield 2000: 27–9, 38–40; 2001: 58, 66; Dreyfus and Dreyfus 2004; Jacobson 2005: 389; Russell 2015a; 2015b: 96; Stichter 2007: 193–4; 2011: 79). Executing the skill analogy is difficult, because skill development, even in well-studied domains like chess that are considerably less fraught than morality, is as yet incompletely understood. In what follows, we articulate some of the promise, and pitfalls, of the approach.

Notice, first, that the skill analogy need not be seen as excluding other accounts of moral improvement; indeed, the analogy might be seen as a rubric available on a variety of approaches. In this section, we'll look at a few recent accounts of moral improvement and how they may be understood under the auspices of skill acquisition.

Virtue ethicists has often invoked (actual or fictitious) virtuous individuals as a source of ethical guidance for those who are not yet virtuous. For instance, Hursthouse (1999: 28) tells us that we may begin to cultivate a virtue-ethical decision process by doing what the virtuous person would do in similar circumstances. If virtue is to be thought of as a skill, then the use of more experienced, or 'skilled', virtuous exemplars to help guide our inexperienced actions is, at least on first glance, quite plausible. Likewise, analogous practices seem to take place in chess: many chess experts have been known to model their own game after another great, as Kasparov did with Alekhine (Kasparov 1996).

Even in the realm of chess, this suggestion has limitations: while it may be a successful technique among chess experts of a high calibre, there is further question whether such modelling is similarly effective for the novice, who is perhaps more prone to errors and to misapplying complicated moves. And however plausible such modelling is in chess, further challenges loom for the moral domain, for the domain of morality is highly complex, and the demands of morality may be highly circumstantial or person-specific. Given this, the practice of consulting a moral exemplar may be both epistemically and practically challenging:

what would that extraordinary person do in this ordinary person's circumstances, and could this ordinary person even do it? Indeed, attempts at emulation may be detrimental, since the novice might be led astray by following the exemplar into challenging moral terrain where the virtuous may stride with assurance: the temperate may incur no risk in dining at a restaurant celebrated for its desserts while they are on a diet, but most of us probably could not.

However, there is another way of thinking about the role of moral exemplars. Recent literature on moral improvement has invoked the use of exemplars for *motivational* purposes—we respond to exceptional excellence in others with admiration (Algoe and Haidt 2009), which inspires and motivates us to imitate that which we admire (Zagzebski 2017: 35). Recently, Engelen, Thomas, Archer and van de Ven (2018) advocated the use of exemplars in moral education, making use of the findings of Rushton and Campbell (1977) that those who observed a role model performing an altruistic action (e.g. donating blood) were more likely to perform that action both immediately afterwards as well as up to six weeks later.

In the domain of academics, role models have shown similar motivational effects on students and their educational outcomes (Lockwood, Jordan, and Kunda 2002; Klopfenstein 2005; Morgenroth, Ryan, and Peters 2015). Here, chess can also offer a suggestive illustration. One explanation for the relatively higher rate of female chess 'dropouts' is the dearth of women chess masters (Chabris and Glickman 2006: 1044; de Bruin, Smits, Rikers, and Schmidt 2008). Given that role models have more positive effects when they are taken to be more relatable (Dijkstra, Kuyper, Buunk, et al. 2008; Han, Kim, Jeong, and Cohen 2017; Lin-Siegler, Ahn, Chen, et al. 2016: 321–3), novice female chess players may lack sufficiently motivating role models. Having a female role model in chess could help inspire another rising female expert to stick with it, just as channelling their 'inner Kipchoge' (the marathon world record-holder) might keep the high-school cross-country runner from falling off that last mile of the race.

What would this motivational aid of exemplars look like in the case of morality? The skill analogy might be reasonably tight in the case of continuing or discontinuing particular moral projects: when the morally fatigued vegetarian falters in the presence of bacon, her admiration for a vegetarian role model may provide the motivational support necessary to stick with vaguely meat-like soy substitutes (*What Would Deborah*

Madison Do?). One important question for the character sceptic concerns how domain-specific this motivational influence is expected to be: is the vegetarian exemplar also likely to be the recycling exemplar? Becoming an expert in morality would require a highly generalized form of expertise, since the reach of morality is plausibly thought to extend over a huge range of human endeavours, across widely varying contexts.

In any event, emulation is likely not to be the whole story, for many skills appear to require instruction and practice; indeed, teachers and coaches need not be exemplars. Many strong chess players receive coaching (de Bruin, Rikers, and Schmidt 2007: 571; de Bruin, Kok, Leppink, and Camp 2014: 19; Gobet and Campitelli 2007: 169), but the contribution of coaching to chess skill is fairly modest (Charness, Krampe, and Mayr 1996; Charness, Tuffiash, Krampe, et al. 2005; Howard 2012). Conversely, practice is known to matter, and matter quite a bit—it has been called 'by far the best predictor of chess rating' (Bilalić, McLeod, and Gobet 2007a: 467). Certainly, the biographies of chess greats indicate that 'intense dedication' is requisite for excellence (Gobet and Campitelli 2007: 161–2). However, we must consider *what kind* of practice will increase proficiency: important factors include receiving feedback and having opportunities for correction of error throughout one's practice (de Bruin, Rikers, and Schmidt 2007: 561; de Bruin, Kok, Leppink, and Camp 2014: 18; Ericsson et al. 1993; Gobet and Campitelli 2007: 160). As for quantity, the famous '10,000 hour rule' is probably a reasonable generalization: while time in practice to become a master varies widely, 10,000 hours is in the vicinity of average (Campitelli and Gobet 2011; Charness et al. 2005).

However, the academic source of the popular 10,000 hour rule, *deliberate practice theory*, notoriously overreaches in claiming that 10,000 hours of serious practice is *sufficient* for expertise (Ericsson, Krampe, and Tesch-Romer 1993: 392; cf. Ericsson, Prietula, and Cokely 2007). You'd expect that other things, like talent, must matter: 10,000 hours is not going to make your average gym rat into LeBron James. And that's what the evidence shows: Hambrick and colleagues' (2014) analysis of six studies found that '[o]n average, deliberate practice explained 34 per cent of the reliable variance in chess performance, leaving 66 per cent unexplained and potentially explainable by other factors' (p. 38). In other areas, practice explained even less: 21 per cent in music, 18 per cent in sports, and only 4 per cent in education (Macnamara, Hambrick, and Oswald 2014: 1615). Clearly, practice isn't the whole story.

Still, practice matters for skill acquisition—the associated effective sizes are pretty robust by psychology standards—so it's probably the best place to start when thinking about moral skill development. And it's certainly intuitive enough: as Aristotle (1984, II, 1099b4–b24) said, obtaining virtue requires 'study and care'.[9] But what would the right sort of practice in morality look like? We can't just pull out our 'morality board' and sit down to practise for several hours a day, keeping all other affairs out of sight and mind. Things get even more complicated with respect to moral learning, since often we don't receive consistent, decisive, or timely feedback.

Given that deliberate and effortful attempts of 'trying harder' to act virtuously often falter as soon we get distracted or become exhausted, many (Besser-Jones 2008: 329; McKenna and Warmke 2017: 728–9; Railton 2011; Stichter 2018: 18–20) have suggested we use the technique of *implementation intentions* to automatize particular desired behaviours in a relatively effortless manner. There is substantial empirical research backing the effectiveness of implementation intentions (Gollwitzer, Brandstätter 1997; Gollwitzer 1999; Gollwitzer and Sheeran 2009), making this is a promising suggestion for virtue cultivation, and one that may fit within the skill-analogy model.[10] Implementation intentions work by cognitively linking a situational cue with a particular behaviour—such as asking for a sparkling water if offered a beer. The situational cue is stored in one's memory, making it more salient and so more readily recognized. Once recognized, the cue automatically triggers the corresponding behaviour. Adopting implementation intentions is one way to instil habits, for this process involves changing conscious intentions into automatic situation-behaviour responses (Achtziger, Bayer, and Gollwitzer 2012).

However, there are substantial trade-offs in relying on implementation intentions. Situational cue or pattern recognition may be highly contextualized and narrow: for example, while expert chess players have superior memory for chess positions, this advantage dissipates for arrangements of pieces that don't make 'chess sense' (Chase and Simon

[9] Whatever the recipe is, it may not be literal study; a series of studies led by Schwitzgebel suggests that professional students of ethics behave no better than anyone else (Schwitzgebel 2009; 2013; Rust and Schwitzgebel 2013).

[10] Stichter (2018) employs the framework of goal automaticity and habit formation in his account of virtue as a skill.

1973).[11] Similar downfalls have been reported with the use of implementation intentions: people who adopted an implementation intention (e.g. 'If I am tempted to drink, then I will call my sponsor') in service of a larger goal (quitting drinking) stuck to their plan of identifying *X* cue and responding with *Y* behaviour, even when a more efficacious path was available (e.g. giving my credit card to a friend while out at the pub) for achieving the same goal (Belyavsky-Bayuk, Janiszewski, and Leboeuf 2010; Parks-Stamm, Gollwitzer, and Oettingen 2007; Masicampo and Baumeister 2012).

Moreover, using implementation intentions has been found to make one *worse* at exhibiting goal-relevant behaviour when the specific implementation intention-invoking situational cue is absent (Bieleke, Legrand, Mignon, and Gollwitzer 2018). One solution might be to adopt more flexible plans or a greater number of implementation intentions, allowing for various situational cues to be accommodated. Yet such modifications have rendered implementation intentions ineffective, for this makes the *if* cue less cognitively accessible (Verhoeven, Adriaanse, de Ridder, et al. 2013) as well as interfering with the strength of *if–then* associations (Vinkers, Adrianse, Kroese, and de Ridder 2015). Such difficulties are considerable enough, we think, to counsel against relying too heavily on implementation intentions for moral improvement.[12]

So far, we've considered whether and to what extent things like instruction, practice, and habit formation matters when it comes to developing expertise. While they do count—to some degree, in at least certain contexts—much is still left unaccounted for. When we consider chess expertise, lots of other factors may matter, but often with small or inconsistent effects: intelligence (Burgoyne et al. 2016: 73), physical fitness (Hinson 2014; Shahade 2015), talent (Howard 2009: 201), personality traits (Bilalić et al. 2007a; 2007b), being left-handed (Campitelli and Gobet 2011: 283–4; Gobet and Campitelli 2007: 168), and even the month of birth (Gobet and Chassy 2008). Additionally, many of these factors have a substantial genetic component, and the extent to which they are 'intervenable' for given individuals may be quite limited (for the genetic

[11] This effect is commonplace, but has not always been found (Bilalić et al. 2007a: 459; Gobet and Simon 2000; Van der Maas and Wagenmaers 2005: 52)

[12] For further discussion on the limitations of using implementation intentions in virtue cultivation, see Waggoner 2021.

component of talent, see Howe, Davidson, and Sloboda 1998: 399–400; for the genetic component of physical fitness, see Bouchard and Rankinen 2001; Mann, Lamberts, and Lambert 2014). In short, many unknowns linger; we are far from giving a comprehensive account of whatever explains the variances in chess skill. And the same is true for the far more expansive and uncertain domain of morality.

We suspect that whatever approximation of a comprehensive theory of moral skill acquisition emerges, it will confirm to the *Lotta-Little Principle* (Doris, forthcoming)*:* typically, many factors are implicated in complex psychological outcomes, and relatively seldom are individual factors implicated especially strongly. With so many factors in play, only seldom will a variable rise to the level of a large effect size. As Ahadi and Diener (1989: 398) put it, 'to expect any psychological variable to correlate with some behavioral criterion on the order of .5 or greater is to deny the complexity of human behavior.' We should expect to find that, whatever the recipe is for moral improvement, the list of ingredients will be large, and few, if any, of the ingredients will have a dominant role in the finished dish.

At present, we know relatively little about what these ingredients are—for example, were some of us born with more 'moral talent'? And what might these talents be? And how—if at all—might these talents be cultivated? If they can't be cultivated, we face the unwelcome implication that some individuals may be barred from the possibility of virtue, just as some will never be able to excel at chess. Furthermore, people may be unlikely to become proficient in *all* of morality, and whatever moral proficiency people attain will likely be rather context-specific and narrow, just as an athlete skilled in one sport will not necessarily—indeed, very seldom—be highly proficient at all sports, or even multiple sports. This is not to say that moral improvement, or even the development of virtue, is impossible. But for any account of moral improvement, our optimism should be bounded: the effect of any particular intervention is likely to be limited, in both magnitude and domain.

32.4 Situationism, agency, and moral responsibility

Thus far, we have focused on disputes about character scepticism and what empirically informed character scepticism entails about the possibility of moral improvement. In this section, we turn to a different but related set of disputes that arose from philosophical reflections on situationist

findings: the nature of human agency and abilities and, in particular, whether practices of holding one another morally responsible are compatible with situationist findings.

It is striking that, apart from a handful of notable exceptions (e.g. Schoeman 1990; Bok 1996; Doris 2002), philosophers interested in responsibility and agency were slow to address the significance of situationist social psychology. At least within the discipline of psychology, most of the attention-grabbing studies in the situationist portfolio, including Milgram's work on obedience in the 1960s and the notorious Stanford Prison Experiment in 1971, were often understood to show something important about freedom and responsibility. By the mid-2000s, partly influenced by debates about character scepticism, a number of philosophers began to contemplate whether the situationist picture entails challenges to standard philosophical accounts of agency and responsibility (Nelkin 2005; Doris and Murphy 2007; Nahmias 2007).

The details of those accounts, and the debates that ensued, were partly shaped by competing approaches within the theory of moral responsibility. So, a few remarks about those background commitments are in order. Putting aside eliminativist views, or views according to which no one is morally responsible (Strawson 1994; Pereboom 2001; Caruso and Morris 2017), most contemporary accounts of responsibility have proceeded from one of two basic pictures about the nature of responsibility: *reason-responsiveness* or *rational capacitarian* accounts (classic examples include Wolf 1990; Fischer and Ravizza 1998), and *identificationist* or *self-expression* accounts (classic examples include Frankfurt 1971; Watson 1975). On both approaches, whether an agent is morally responsible for some action depends on whether the action stands in the right relationship to a distinctive feature of the agent. The rational capacitarian holds that non-derivative responsibility for some behaviour requires that it be rooted in some rational process, faculty, or mechanism. (On a given approach, there might be further requirements above and beyond the minimal requirement of mediation by some rational element; rational capacitarians can also hold that culpable actions must be voluntary, or that they manifest a certain quality of will. For ease of exposition, we'll ignore these complexities.)

The identificationist locates an agent's responsibility for some behaviour in the coherence of that behaviour with some privileged psychological attitude or complex of attitudes, such as higher-order desires

(desires about desires) or valuings. Here too, one might add further conditions on the minimally necessary condition. There are a variety of views that don't neatly fit into either of these families (e.g. Scanlon 2008), or that explicitly require some further addition of distinctive agential powers, including emergent or indeterministic causal powers (Clarke and Capes 2017; Kane 1996). Moreover, some contemporary accounts can be plausibly characterized in different ways (Vargas 2020: 411–18). However, it was within 'the big two' approaches that debates about situationism unfolded.

In drawing a distinction between these families of philosophical approaches, it would be a mistake to presume that there was little in common between them. On all sides, recent philosophical work has tended to approach responsibility in large part as a natural, psychological, and social phenomenon, frequently characterized in terms of the moral psychology of responsibility practices (Shoemaker 2015; Nelkin and Pereboom, forthcoming; see related chapters in this volume, including Chapters 27 and 35). Indeed, it is this shared commitment that propelled theorists of otherwise notably different convictions to take seriously challenges generated by situationist findings. That said, at least at the outset, the details of the debates about situationism and responsibility tended to unfold in forms specific to the two main theoretical approaches.

32.4.1 Reasons-responsiveness

Situationism has sometimes been thought to bear on moral responsibility via some context-specific impairment to normative or rational competence (Doris 2002: 138; Doris and Murphy 2007). Here's the thought: according to reasons-responsiveness or rational capacitarian theories, to be responsible an agent has to be able to recognize and respond to relevant normative reasons. What situationist findings seem to show is that in a wide range of cases, agents fail to recognize and respond to normatively relevant considerations, and indeed, they often respond to normatively or rationally irrelevant features of the practical context (Nelkin 2005). To the extent to which this happens, agents seem to lack the capacity or ability to respond to relevant normative considerations.

The situationist threat dovetails with a family of related claims, common to some early 2000s neuroscience, cognitive, and social psychology, according to which the vast majority of human behaviour is automatic and non-conscious (Bargh and Ferguson 2000; Wegner 2002; for

discussion see Mele 2009; Nahmias 2010; Vargas 2013b; Doris 2015). Jointly, these findings seem to entail that agents are at least often, and maybe usually, unaware of the basis of their actions, that those actions are frequently propelled by irrelevant features of the context, and that agents are widely self-deceived about the foregoing.

Rational capacitarians have offered two distinct paths of response to this family of concerns: *accommodation* and *resistance*. Echoing moves made by some apologists for revisionist conceptions of virtue (see Section 32.2), the path of accommodation allows that situations can impair an agent's normative competence, but insists that a suitably nuanced picture of rational abilities can allow for highly localized impairments of capacity and/or diminished abilities so that responsibility practices can continue in roughly their current forms, albeit with diminished frequency and/or degrees of responsibility (Vargas 2013b). On this approach, although situationism has implications for responsibility, it is less a matter of situationist findings undermining responsibility practices as a whole than a matter of attenuating the frequency or degree with which we hold people responsible. If one holds that the responsibility-relevant powers of agents are partly ecological, or a matter of non-intrinsic features of agents such as opportunities or circumstances (as in Vargas 2013a; Brink and Nelkin 2013; Washington and Kelly 2016; Chapter 27 in this volume), then the import of situationism for rational capacities may be primarily a matter of its highlighting the fragile nature of the ecological conditions required for responsibility-relevant abilities. On this approach, even if we are tempted to 'speak with the folk' in thinking that people always have a general capacity to deliberate about etiquette, the theorist's responsibility-relevant capacity will be something narrower, e.g. being awake, not subject to various kinds of distractions, and otherwise free of rational-disrupting situational effects. For the accomodationist rational capacitarian, the upshot of situationism is that we must forfeit reliance on a robustly cross-situationally stable notion of a general capacity in favour of a more fine-grained and contextual picture of capacities.

The path of resistance rejects the situationist threat to responsibility as premature, or at least overstated. The crucial idea is to distinguish between the possession and exercise of a capacity (Brink 2013; Vargas 2013b; see also Fischer 2018: 251–2 on 'good enough' rational capacities). On this approach, situationist evidence gives us reason to think that situational factors affect whether agents exercise their responsibility-

relevant capacities (or their *rational abilities*, as it is sometimes put). However, diminished exercises of a capacity are compatible with ongoing presence of the capacity. Since suitability for responsibility assessments only requires that the considered agent possess the responsibility-relevant capacity, and not that they have correctly exercised it, situationist findings that report changes in behaviour do not by themselves show changes in the underlying abilities required of responsible agents.

One can mix and match elements of these responses, insisting that rational abilities often persist in the face of situationist pressures, and that the evidence does not yet show the absence of rational capacities (cf. McKenna and Warmke 2017: 719) while also allowing that there may be times when those pressures alter the responsibility-relevant abilities, whether directly or via impairment or improvement of the ecological conditions on responsibility (Vargas 2013b: 343).

An important development in the wake of these ruminations on the situationist challenge to responsibility has been a renewed appreciation of the difficulty of spelling out the responsibility-relevant notion of ability. The basic issue is not new. Disputes about the conditional analysis of 'can' (see Kane 1996: 4758) and more recent efforts by 'the new dispositionalists' have a substantial literature around them (see Clarke 2009; Franklin 2018). Still, it was the rational capacitarian responses to situationism that led Carolina Sartorio to press what she calls the *demarcation challenge*, which concerns how we are to go about identifying and constraining the features of context that matter in the assessment of an agent's responsibility-relevant ability:

> we need some principled reason to single out the aspects of the actual circumstances that we can vary from the aspects of the circumstances that we must held fixed in order to assess an agent's reasons-responsiveness on a certain occasion [. . . we] need to say more about which [worlds] are relevant and which ones aren't. (Sartorio 2018: 800)

In the context of motivating her own alternative to rational capacitarianism, Sartorio asserts that (apart from McKenna, 2005) very little has been said in the literature about the demarcation problem, despite its deserving immediate attention for rational capacitarians (2018: 800–801). There is some reason to resist that assessment.

For example, at least within instrumentalist accounts of responsibility—the family of views that emphasize the importance of

instrumental considerations in especially the justification of responsibility practices—there was already something of a literature that anticipated and sought to address the demarcation problem. Vargas (2013a: 217–33, esp. 217–22; 2017: 234–6; 2018) argues that we can capture the modal features that matter for responsibility in terms of a capacity constructed from the instrumentalist considerations that justify responsibility practices. On that account, the relevant situational features (or counterfactuals) are fixed by constraints about existing norms and psychological dispositions and by what construction of ability would best enable agents to recognize and respond to moral considerations in the actual world. That approach has been critiqued and further developed by McGeer (2015: 2645–8) and taken up in a different way by McGeer and Pettit (2015). Although there are important differences among these accounts, they all share the idea that instrumental features structuring the practice of responsibility—in particular, the utility of the practice in cultivating a desirable form of agency—is largely determinative in specifying the relevant counterfactuals.

Brink's (2013) discussion of situationism and reasons-responsiveness provides a different approach to the demarcation problem. Brink (2013: 141) anticipates Sartorio's observation that there is some complexity in how to constrain the relevant counterfactuals. He goes on to suggest a number of plausible constraints on those counterfactuals, including: constraints of familiarity; a metaphysical restriction of counterfactuals to the specific agent, coupled with a practical allowance that we may try to assess this by looking to other agents in the same context; the possibility of performance mistakes and errors; and the requirement that counterfactual cases are cases where there is regular performance of the action (p. 141). He expects that 'appropriate counterfactual evidence would vindicate ordinary assumptions people have' about rational capacities (p. 142), and he employs reflections from the criminal law to show the relatively recognizable and powerful ways everyday distinctions can be deployed. So, we might think of this as a 'common-sense' approach to abilities, one that allows for some difficult and indeterminate cases, while insisting that we generally have a reasonably good grip on which differences matter for actual and possible abilities.

Whatever the merits of these approaches may be, situationism has undoubtedly been a spur to important developments in rational capacitarian approaches to moral responsibility.

32.4.2 Self-expression

Identificationist (or self-expression) views typically maintain that agents are morally responsible for some behaviour to the extent that the behaviour expresses or meshes with authoritative or privileged aspects of the agent's psychology. This is sometimes put in terms of expressing or meshing with the agent's 'real' or 'deep' self, or in terms of coherence with the agent's evaluative commitments or endorsements. These differences can, of course, substantively affect the details of the theory. Even so, traditional versions of these views tend to share the idea that there is a kernel or normatively privileged psychological nugget, such that behaviour standing in the right relationship with that nugget is behaviour for which the agent is morally responsible.

One threat to such views, considered by Doris (2007), is that such nuggets look remarkably close to the character or cross-situationally stable dispositions appealed to in virtue theories. So, although self-expression views may avoid worries about whether agent actions reliably flow through rational capacities, it is not obvious that the accounts don't face comparable or greater worries about whether there is a stable nugget or kernel of normatively privileged psychology that suffices to constitute a deep self. If an agent's commitments are either indeterminate because of their situational fragility, or else so finely granulated that one has only a perspective-in-this-circumstance, it is unclear how much confidence we should have in our assessments of responsibility.

A different line of concern emerges if one thinks that self-expression views require that the agent be consciously aware of the attitude being expressed (Levy 2014: 87–108), or alternatively, that responsible agency requires that judgments and behaviour be ordered by accurate self-conscious reflection (Doris 2015: 17–40). Views committed to such pictures of responsible agency take on board commitments that seem undermined by situationist findings and other work in contemporary cognitive science. Although we may sometimes know our motives and be aware of the values we are expressing, the empirical work suggests that we are often confabulating, self-deceived, or simply unaware of the causes of our behaviour and choices.

Traditional self-expression views are backward-looking, in the sense that the central theoretical question in the assessment of responsibility is in terms of a past or current bit of behaviour with the agent's existing nugget. However, Doris (2015) has developed an alternative

that makes the import of coherence with the nugget a forward-looking one––and indeed, less about coherence with an existing nugget than the construction of a socially outsourced nugget. On his 'collaborativist' account, responsibility is grounded in coherence of conduct with an agent's desires and values. However, the function of that coherence is less about reflecting or expressing an antecedently existing or conscious self, as in backward-looking views. Rather, the function is primarily forward-looking, about the agent binding herself to explaining, justifying, and being called to account on the basis of those values. Thus, the agent's values can be discovered, and indeed created, in the process of (frequently) social and collaborative reasoning about action and its significance.

In response to Doris's account, a number of authors have pressed the concern that values are vulnerable to the same situationist pressures that arise for virtues, namely, that they lack sufficient cross-situational stability (Arpaly 2018: 755; Nelkin 2018: 271–2; Vargas 2018: 265). However, as Doris (2018) has noted, one important feature of values is that we do not expect them to have as robust behavioural consistency as has been sometimes imputed to virtuous action. One can value fitness even while failing to make good on that value, and akratic action is an apparently familiar phenomenon. So, for Doris's forward-looking valuational account of responsibility, the problem of a potentially implausible nugget of cross-situationally stable psychological dispositions is diminished once we allow that an agent might value something without those values manifesting in behaviour. Responsibility practices retain their efficacy by pressuring individuals to justify choices and interpret their own behaviour in light of the asserted (or sometimes dialogically discovered) values expressed by those agents. As with the accounts emphasizing 'inner' virtue already discussed, there are questions about whether the valuing-relevant psychological states will be unduly subject to situational perturbances; on Doris' (2015) collaborativist account of agency, the answer is supposed to lie––as it does in socially sustained accounts of virtue (and, for that matter, rational capacities)––in the support of facilitating exterior scaffolding.

32.4.3 Further developments

It is worth noting a striking parallel in how valuational and rational capacitarian theories handle the problem of behaviour at odds with the feature of agency that grounds responsibility. In response to situationist pressures, responsibility theorists of different persuasions have drawn a

distinction between possession of a property and its behavioural manifestation. The details differ, but the shared insight is that possession of the responsibility-making feature (be it values or rational capacities) is compatible with failures to manifest that property in behaviour. This buys theorists of either stripe a certain degree of wiggle room in accommodating behavioural findings.

One might suppose that the import of social psychological findings for moral responsibility ends there. However, there is reason to doubt that responsibility theorists can sit tight. For example, Rudy-Hiller (2020) has recently argued that the principal import of social psychological findings for the theory of moral responsibility is not that it shows we lack some or another responsibility-enabling feature of agency. Instead, it highlights how difficult it is for us to be morally responsible. Responsibility might be like some virtue theorists insist the virtuous person is—a rare achievement. In a different vein, Piovarchy (forthcoming) has argued that situationist findings suggest that many of us are not consistently and fully committed to the moral values that ground our complaints about others. If so, then we may frequently lack the standing to blame others for their wrongdoing. These recent developments suggest that philosophers are not yet done mining situationist social psychology for philosophical insight about moral agency and responsibility.

32.5 Conclusion

The virtue ethics–situationism debate dates to the beginnings of moral psychology as a robustly interdisciplinary field, joining philosophy with the human sciences and beyond. Indeed, the abiding interest of the debate is likely an important factor in vivifying moral psychology as an academic discipline And—like the field of moral psychology more broadly—the debate has expanded far beyond its beginnings: no longer a narrowly focused critique of virtue ethics in philosophy, reflections on moral agents in light of empirical research now spans the academy, drawing researchers with both theoretical and empirical proclivities from a wide variety of fields (e.g. Miller, Furr, Knobel, and Fleeson 2015). And just as the wider field of moral psychology is (as the chapters in this Handbook testify) vibrantly flourishing, the debate over character scepticism continues to uncover new avenues of progress in understanding moral personality.

References

Achtziger, A., U. Bayer, and B. Gollwitzer. 2012. Committing to implementation intentions: attention and memory effects for selected situational cues. Motivation and Emotion 36: 287–300.

Adams, R. M. 2006. A Theory of Virtue: Excellence in Being for the Good. Oxford: Oxford University Press.

Ahadi, S., and E. Diener. 1989. Multiple determinants and effect size. Journal of Personality and Social Psychology 56(3): 398–406.

Alfano, M. 2013. Character as Moral Fiction. Cambridge: Cambridge University Press.

Alfano, M. 2018. A plague on both your houses: virtue theory after situationism and repligate. Teoria 38(2): 115–22.

Algoe, S. B., and J. Haidt. 2009. Witnessing excellence in action: the 'other-praising' emotions of elevation, gratitude, and admiration. Journal of Positive Psychology 4(2): 105–27.

Alston, W. P. 1975. Traits, consistency and conceptual alternatives for personality theory. Journal for the Theory of Social Behaviour 5(1): 17–48.

Anderson, C. A. 2001. Heat and violence. Current Directions in Psychological Science 10(1): 33–8.

Annas, J. 2003. Virtue ethics and social psychology. A Priori 2: 20–33.

Annas, J. 2011. Intelligent Virtue. Oxford: Oxford University Press.

Anscombe, G. E. M. 1958. Modern moral philosophy. Philosophy 33(124): 1–19.

Ariely, D., and K. Wertenbroch. 2002. Procrastination, deadlines, and performance: self-control by precommitment. Psychological Science 13: 219–24.

Arpaly, N. 2018. Comments on Talking to Our Selves. Philosophy and Phenomenological Research 97(3): 753–7.

Aristotle. 1984. The Complete Works of Aristotle, ed. J. Barnes. Princeton, NJ: Princeton University Press.

Bargh, J. A., and M. J. Ferguson. 2000. Beyond behaviorism: on the automaticity of higher mental processes. Psychological Bulletin 126(6): 925–45.

Baron, J. 1994. Nonconsequentialist decisions. Behavioral and Brain Sciences 17: 1–42.

Baron, J. 2001. Thinking and Deciding, 3rd edn. Cambridge: Cambridge University Press.

Baron, R. A. 1997. The sweet smell of . . . helping: effects of pleasant ambient fragrance on prosocial behavior in shopping malls. Personality and Social Psychology Bulletin 23(5): 498–503.

Belyavsky-Bayuk, J., C. Janiszewski, and R. A. Leboeuf. 2010. Letting good opportunities pass us by: examining the role of mind-set during goal pursuit. Journal of Consumer Research 37: 570–83.

Besser-Jones, L. 2008. Social psychology, moral character, and moral fallibility. Philosophy and Phenomenological Research 76: 310–32.

Bieleke, M., E. Legrand, A. Mignon, and P. M. Gollwitzer. 2018. More than planned: implementation intention effects in nonplanned situations. Acta Psychologica 184: 64–74.

Bilalić, M., : McLeod, and F. Gobet. 2007a. Does chess need intelligence? A study with young chess players. Intelligence 35(5): 457–70.

Bilalić′, M.: McLeod, and F. Gobet. 2007b. Personality profiles of young chess players. Personality and Individual Differences 42(6): 901–10.

Bloomfield: 2000. Virtue epistemology and the epistemology of virtue. Philosophy and Phenomenological Research 60(1): 23–43.

Bloomfield: 2001. Moral Reality. Oxford: Oxford University Press.

Bloomfield: 2014. Some intellectual aspects of the cardinal virtues. In Oxford Studies in Normative Ethics, vol. 3, ed. M. Timmons. Oxford: Oxford University Press. pp. 287-313.

Bok, H. 1996. Acting without choosing. Noûs 30(2): 174–96.

Bouchard, C., and T. Rankinen. 2001. Individual differences in response to regular physical activity. Medicine and Science in Sports and Exercise 3: S446–51.

Brink, D. O. 2013. Situationism, responsibility, and fair opportunity. Social Philosophy and Policy 30: 121–49.

Brink, D. O., and D. Nelkin. 2013. Fairness and the architecture of responsibility. Oxford Studies in Agency and Responsibility 1: 284–314.

Burgoyne, A. P., G. Sala, F. Gobet, B. N. Macnamara, G. Campitelli, and D. Z. Hambrick. 2016. The relationship between cognitive ability and chess skill: a comprehensive meta-analysis. Intelligence 59: 72–83.

Campitelli, G., and F. Gobet. 2011. Deliberate practice necessary but not sufficient. Current Directions in Psychological Science 20(5): 280–85.

Caruso, G. D., and S. G. Morris. 2017. Compatibilism and retributive desert moral responsibility: On what is of central philosophical and practical importance. Erkenntnis 82(4): 837–55

Cawley, J., and J. A. Price. 2011. Outcomes in a program that offers financial rewards for weight loss. In Economic Aspects of Obesity. Washington, DC: National Bureau of Economic Research. pp. 91-126.

Ceci, S. J. 1993a. Teaching for transfer: the 'now-you-see-it-now-you-don't' quality of intelligence in context. In The Edyth Bush Symposium on Intelligence, ed. H. Rosselli. Orlando, FL: Academic Press.

Ceci, S. J. 1993b. Contextual trends in intellectual development. Developmental Review 13(4): 403–35.

Chabris, C. F., and M. E. Glickman. 2006. Sex differences in intellectual performance: analysis of a large cohort of competitive chess players. Psychological Science 17(12): 1040–46.

Chambers, C. 2017. The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice. Princeton, NJ: Princeton University Press.

Charness, G., and M. Dufwenberg. 2006. Promises and partnership. Econometrica 74(6): 1579–1601.

Charness, N., R. T. Krampe, and U. Mayr. 1996. The role of practice and coaching in entrepreneurial skill domains: an international comparison of life-span chess skill acquisition. In The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports and Games, ed. K. A. Ericsson. Mahwah, NJ: Erlbaum. pp. 51-80.

Charness, N., M. Tuffiash, R. Krampe, E. Reingold, and E. Vasyukova. 2005. The role of deliberate practice in chess expertise. Applied Cognitive Psychology 19(2): 151–65.

Chase, W. G., and H. A. Simon. 1973. Perception in chess. Cognitive Psychology 4(1): 55–81.

Ciurria, M. 2014. Answering the situationist challenge: a defense of virtue ethics as preferable to other ethical theories. Dialogue 53(4): 651–70.

Clarke, R. 2009. Dispositions, abilities to act, and free will: the new dispositionalism. Mind 118(470): 323–51.

Clarke, R., and J. Capes. 2017. Incompatibilist (nondeterministic) theories of free will. In Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta. https://plato.stanford.edu/archives/spr2017/entries/incompatibilism-theories/>

Cohen, R., and J. D'Cruz. 2016. Promises and consistency. In Questions of Character, ed. I. Fileva. New York: Oxford University Press.

Darley, J. M., and C. D. Batson. 1973. 'From Jerusalem to Jericho': a study of situational and dispositional variables in helping behavior. Journal of Personality and Social Psychology 27(1): 100–108.

Darley, J. M., and Latane, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. Journal of Personality and Social Psychology 8(4): 377–83.

de Bruin, A. B., E. M. Kok, J. Leppink, and G. Camp. 2014. Practice, intelligence, and enjoyment in novice chess players: a prospective study at the earliest stage of a chess career. Intelligence 45: 18–25.

de Bruin, A. B., R. M. J. P. Rikers, and H. G. Schmidt. 2007. The influence of achievement motivation and chess-specific motivation on deliberate practice. Journal of Sport and Exercise Psychology 29(5): 561–83.

de Bruin, A. B., N. Smits, R. M. J. P. Rikers, and H. G. Schmidt. 2008. Deliberate practice predicts performance over time in adolescent chess players and drop-outs: a linear mixed models analysis. British Journal of Psychology 99(4): 473–97.

Dijkstra, P., H. Kuyper, A. P. Buunk, G. van der Werf, and Y. van der Zee. 2008. Social comparison in the classroom: a review. Review of Educational Research 78(4): 828–79.

Doris, J. M. 1998. Persons, situations, and virtue ethics. Noûs 32(4): 504–30.

Doris, J. M. 2002. Lack of Character: Personality and Moral Behavior. New York: Cambridge University Press.

Doris, J. M. 2005. Replies: evidence and sensibility. Philosophy and Phenomenological Research 71(3): 656–77.

Doris, J. M. 2007. Out of character: on the psychology of excuses in the criminal law. In Ethics in Practice, 3rd edn, ed. H. Lafolette. Malden, MA: Blackwell.

Doris, J. M. 2010. Heated agreement: lack of character as being for the good. Philosophical Studies 148(1): 135–46.

Doris, J. M. 2015. Talking to Our Selves: Reflection, Ignorance, and Agency. Oxford: Oxford University Press.

Doris, J. M. 2018. Making do without (reflection): a (very partial) response to Arpaly, Tiberius, and Kane. Philosophy and Phenomenological Research 97(3): 771–90.

Doris, J. M. Forthcoming. Character Trouble: Undisciplined Essays on Personality and Agency. Oxford: Oxford University Press.

Doris, J. M., and D. Murphy. 2007. From My Lai to Abu Ghraib: the moral psychology of atrocity. Midwest Studies in Philosophy 31: 25–55.

Doris, J. M., and S. Stich. 2005. As a matter of fact: empirical perspectives on ethics. In The Oxford Handbook of Contemporary Philosophy, ed. F. Jackson and M. Smith. Oxford: Oxford University Press. pp. 114-152.

Dreyfus, H. L., and S. E. Dreyfus. 2004. The ethical implications of the five-stage skill-acquisition model. Bulletin of Science, Technology and Society 24(3): 251–64.

Engelen, B., A. Thomas, A. Archer, and N. van de Ven. 2018. Exemplars and nudges: combining two strategies for moral education. Journal of Moral Education 47(3): 346–65.

Ericsson, K. A., R. T. Krampe, and C. Tesch-Römer. 1993. The role of deliberate practice in the acquisition of expert performance. Psychological Review 100(3): 363–406.

Ericsson, K. A., M. J. Prietula, and E. T. Cokely. 2007. The making of an expert. Harvard Business Review 85: 114–21.

Fischer, J. M. 2018. On John Doris's Talking to Our Selves. Social Theory and Practice 44(2): 247–53.

Fischer, P., J. I. Krueger, T., Greitemeyer, et al. 2011. The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. Psychological Bulletin 137(4): 517–37.

Fischer, J. M., and M. Ravizza. 1998. Responsibility and Control: A Theory of Moral Responsibility. New York: Cambridge University Press.

Flanagan, O. 1991. Varieties of Moral Personality: Ethics and Psychological Realism. Cambridge, MA: Harvard University Press.

Fleeson, W., and R. M. Furr. 2016. Do broad character traits exist? Repeated assessments of individuals, not group summaries from classic experiments, provide the relevant evidence. In Questions of Character, ed. I. Fileva. New York: Oxford University Press. pp. 231-248.

Foot: 1978. Virtues and Vices: And Other Essays in Moral Philosophy. Berkeley: University of California Press.

Frankfurt, H. 1971. Freedom of the will and the concept of a person. Journal of Philosophy 68(1): 5–20.

Franklin, C. E. 2018. A Minimal Libertarianism: Free Will and the Promise of Reduction. New York: Oxford University Press.

Fridland, E. 2017. Motor skill and moral virtue. Royal Institute of Philosophy Supplements 80: 139–70.

Funder, D. C., and D. J. Ozer. 2019. Evaluating effect size in psychological research: sense and nonsense. Advances in Methods and Practices in Psychological Science 2: 156–68.

Gilovich, T., T. Griffin, and D. Kahneman. 2002. Heuristics and Biases: The Psychology of Intuitive Judgment. New York: Cambridge University Press.

Gobet, F., and G. Campitelli. 2007. The role of domain-specific practice, handedness, and starting age in chess. Developmental Psychology 43(1): 159–72.

Gobet, F., and P. Chassy. 2008. Season of birth and chess expertise. Journal of Biosocial Science 40(2): 313–16.

Gobet, F., and H. A. Simon. 2000. Five seconds or sixty? Presentation time in expert memory. Cognitive Science 24(4): 651–82.

Gollwitzer: W. 1999. Implementation intentions: strong effects of simple plans. American Psychologist 54(7): 493–503.

Gollwitzer: W., and V. Brandstätter. 1997. Implementation intentions and effective goal pursuit. Journal of Personality and Social Psychology 73: 186–99.

Gollwitzer: M., and P. Sheeran. 2009. Self-regulation of consumer decision making and behavior: the role of implementation intentions. Journal of Consumer Psychology 19: 593–607.

Grover, L. 2012. The evaluative integration of local character traits. Journal of Value Inquiry 46: 25–37.

Hambrick, D. Z., F. L. Oswald, E. M. Altmann, E. J. Meinz, F. Gobet, and G. Campitelli. 2014. Deliberate practice: Is that all it takes to become an expert? Intelligence 45: 34–45.

Han, H., Kim, J., Jeong, C., and Cohen, G. L. (2017). Attainable and relevant moral exemplars are more effective than extraordinary exemplars in promoting voluntary service engagement. Frontiers in Psychology 8: 283.

Harman, G. 1999. Moral philosophy meets social psychology: virtue ethics and the fundamental attribution error. Proceedings of the Aristotelian Society 99: 315–31.

Harman, G. 2000. The nonexistence of character traits. Proceedings of the Aristotelian Society 100: 223–6.

Harman, G. 2001. Virtue ethics without character traits. In Fact and Value, ed. A. Byrne, R. C. Stalnaker, and R. Wedgewood. Cambridge, MA: MIT Press

Harman, G. 2003. No character or personality. Business Ethics Quarterly 13(1): 87–94.

Harman, G. 2009. Skepticism about character traits. Journal of Ethics 13(2–3): 235–42.

Hinson, M. 2014. Chexercise: chess and physical fitness. Daily Princetonian, 19 Oct. http://dailyprincetonian.com/sports/2014/10/chexercise-chess-and-physical-fitness/

Hochstein, E. 2017. When does 'folk psychology' count as folk psychological? British Journal for the Philosophy of Science 68: 1125–47.

Howard, R. W. 2009. Individual differences in expertise development over decades in a complex intellectual domain. Memory and Cognition 37(2): 194–209.

Howard, R. W. 2012. Longitudinal effects of different types of practice on the development of chess expertise. Applied Cognitive Psychology 26(3): 359–69.

Howe, M. J., J. W. Davidson, and J. A. Sloboda. 1998. Innate talents: reality or myth? Behavioral and Brain Sciences 21(3): 399–407.

Hursthouse, R. 1999. On Virtue Ethics. Oxford: Oxford University Press.

Hutton, E. L. 2006. Character, situationism, and early Confucian thought. Philosophical Studies 127(1): 37–58.

Isen, A. M., and P. F. Levin. 1972. Effect of feeling good on helping: cookies and kindness. Journal of Personality and Social Psychology 21(3): 384–8.

Jacobson, D. 2005. Seeing by feeling: virtues, skills, and moral perception. Ethical Theory and Moral Practice 8(4): 387–409.

Kahneman, D.: Slovic, and A. Tversky. 1982. Judgment Under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press.

Kahneman, D., and A. Tversky. 1982. The simulation heuristic. In Judgment Under Uncertainty: Heuristics and Biases, ed. D. Khaneman: Slovic, and A. Tversky. Cambridge: Cambridge University Press.

Kamtekar, R. 2004. Situationism and virtue ethics on the content of our character. Ethics 114: 458–91.

Kane, R. 1996. The Significance of Free Will. Oxford: Oxford University Press.

Kanngiesser, P., J. Sunderarajan, and J. K. Woike. 2020. Keeping them honest: promises reduce cheating in adolescents. Behavioral Decision Making 34(2): 183–98.

Kasparov, G. 1996. Foreword. In Alexander Alekhine's Best Games. New York: Henry Holt .

Kenrick, D. T., and S. W. MacFarlane. 1986. Ambient temperature and horn honking: a field study of the heat/aggression relationship. Environment and Behavior 18(2): 179–91.

Kleingeld: 2015. Consistent egoists and situation managers: two problems for situationism. Philosophical Explorations 18(3): 344–61.

Klopfenstein, K. 2005. Beyond test scores: the impact of black teacher role models on rigorous math taking. Contemporary Economic Policy 23(3): 416–28.

Kruger, J., and D. Dunning. 1999. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. Journal of Personality and Social Psychology 77: 1121–34.

Kurth, C. 2021. Cultivating disgust: prospects and moral implications. Emotion Review 13(2): 101-112.

Latané, B., and S. Nida. 1981. Ten years of research on group size and helping. Psychological Bulletin 89: 308–24.

Levy, N. 2014. Consciousness and Moral Responsibility. New York: Oxford University Press.

Lin-Siegler, X., J. N. Ahn, J. Chen, F. F. A. Fang, and M. Luna-Lucero. 2016. Even Einstein struggled: effects of learning about great scientists' struggles on high school students' motivation to learn science. Journal of Educational Psychology 108(3): 314–28.

Lockwood, P., C. H. Jordan, and Z. Kunda. 2002. Motivation by positive or negative role models: regulatory focus determines who will best inspire us. Journal of Personality and Social Psychology 83: 854–64.

Lott, M. 2014. Situationism, skill, and the rarity of virtue. Journal of Value Inquiry 48(3): 387–401.

Machery, E. 2010. The bleak implications of moral psychology. Neuroethics 3(3): 223–31.

Macnamara, B. N., D. Z. Hambrick, and F. L. Oswald. 2014. Deliberate practice and performance in music, games, sports, education, and professions a meta-analysis. Psychological Science 25(8): 1608–18.

Magundayao, J. A. M. 2013. Dispositions and skills: an argument for virtue ethics against situationism. Kritike 7(1): 96–114.

Mann, T. N., R. P. Lamberts, and M. I. Lambert. 2014. High responders and low responders: factors associated with individual variation in response to standardized training. Sports Medicine 44(8): 1113–24.

Masicampo, E. J., and R. F. Baumeister. 2012. Committed but close-minded: when making a specific plan for a goal hinders success. Social Cognition 30: 37–55.

Mathews, K. E., and L. K. Canon. 1975. Environmental noise level as a determinant of helping behavior. Journal of Personality and Social Psychology 32(4): 571–7.

McGeer, V. 2012. Co-reactive attitudes and the making of moral community. In Emotions, Imagination, and Moral Reasoning, ed. R. Langdon and C. Mackenzie. New York: Psychology Press.

McGeer, V. 2015. Building a better theory of responsibility. Philosophical Studies 172(10): 2635–49.

McGeer, V., and P. Pettit. 2015. The hard problem of responsibility. In Oxford Studies in Agency and Responsibility, vol. 3, ed. D. Shoemaker. Oxford: Oxford University Press. 160-188.

McKenna, M. 2005. The relationship between autonomous and morally responsible agency. In Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy, ed. J. S. Taylor. Cambridge: Cambridge University Press. pp. 205-234.

McKenna, M., and B. Warmke. 2017. Does situationism threaten free will and moral responsibility? Journal of Moral Philosophy 14: 698–733.

Mehl, M. R., K. L. Bollich, J. M. Doris, and S. Vazire. 2015. Character and coherence: testing the stability of naturalistically observed daily moral behavior. In

Character: New Directions from Philosophy, Psychology, and Theology, ed. C. Miller, M. R. Furr, A. Knobel, and W. Fleeson. Oxford: Oxford University Press. pp. 630-651.

Mele, A. 2009. Effective Intentions: The Power of the Conscious Will. New York: Oxford University Press.

Merritt, M. W. 2000. Virtue ethics and situationist personality psychology. Ethical Theory and Moral Practice 3(4): 365–83.

Merritt, M., J. Doris, and G. Harman. 2010. Character. In The Moral Psychology Handbook, ed. J. Doris. Oxford: Oxford University Press. pp. 355-401.

Milgram, S. 1974. Obedience to Authority: An Experimental View. New York: Harper & Row.

Miller, C. B. 2003. Social psychology and virtue ethics. Journal of Ethics 7(4): 365–92.

Miller, C. B. 2013. Moral Character: An Empirical Theory. Oxford: Oxford University Press.

Miller, C. B. 2014. Character and Moral Psychology. Oxford: Oxford University Press.

Miller, C. B. 2017. Character and situationism: new directions. Ethical Theory and Moral Practice 20: 459–71.

Miller, C. B., R. M. Furr, A. Knobel, and W. Fleeson (eds) 2015. Character: New Directions from Philosophy, Psychology, and Theology. Oxford: Oxford University Press.

Mischel, W. 1968. Personality and Assessment. New York: Wiley

Morgenroth, T., M. K. Ryan, and K. Peters. 2015. The motivational theory of role modeling: how role models influence role aspirants' goals. Review of General Psychology 19(4): 465–83.

Mower, D. S. 2013. Situationism and Confucian virtue ethics. Ethical Theory and Moral Practice 16(1): 113–37.

Olin, L., and J. M. Doris. 2014. Vicious minds: virtue epistemology, cognition, and skepticism. Philosophical Studies 168(3): 665–92.

Nahmias, E. 2007. Autonomous agency and social psychology. In Cartographies of the Mind: Philosophy and Psychology in Intersection, ed. M. Marraffa, M. Caro, and F. Ferretti. Dordrecht: Springer. pp. 169-185.

Nahmias, E. 2010. Scientific challenges to free will. In A Companion to the Philosophy of Action, ed. T. O'Connor and C. Sandis. Malden, MA: Wiley-Blackwell. pp. 345-356.

Nelkin, D. 2005. Freedom, responsibility, and the challenge of situationism. Midwest Studies in Philosophy 29(1): 181–206.

Nelkin, D. 2018. Responsibility and ignorance of the self. Social Theory and Practice 44(2): 267–78.

Nelkin, D., and D. Pereboom (eds) Forthcoming. The Oxford Handbook of Moral Responsibility. New York: Oxford University Press.

Nisbett, R. E., and E. Borgida. 1975. Attribution and the psychology of prediction. Journal of Personality and Social Psychology 32: 932–43.

Nisbett, R. E., and L. Ross. 1980. Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice Hall

Papish, L. 2017. CAPS psychology and the empirical adequacy of Aristotelian virtue ethics. Ethical Theory and Moral Practice 20(3): 537–49.

Parks-Stamm, E. K.: M. Gollwitzer, and G. Oettingen. 2007. Action control by implementation intentions: effective cue detection and efficient response initiation. Social Cognition 25: 247–64.

Pereboom, D. 2001. Living Without Free Will. Cambridge: Cambridge University Press.

Pettit: 2015. The Robust Demands of the Good: Ethics with Attachment, Virtue, and Respect. Oxford: Oxford University Press.

Piovarchy, A. Forthcoming. Situationism, subjunctive hypocrisy and standing to blame. Inquiry: Critical Thinking Across the Disciplines.

Railton, P. 2004. Towards an ethics that inhabits the world. In The Future for Philosophy, ed. B. Leiter. Oxford: Clarendon Press. pp. 265-284.

Railton, P. 2011. Two cheers for virtue: Or, might virtue be habit forming? In Oxford Studies in Normative Ethics, vol. 1, ed. M. Timmons. New York: Oxford University Press. pp. 295-330.

Roberts, B. W., N. R. Kuncel, R. L. Shiner, A. Caspi, and L. R. Goldberg. 2007. The power of personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. Perspectives on Psychological Science 2(4): 313–45.

Roberts, R. C. 2013. Emotions in the Moral Life. Cambridge: Cambridge University Press.

Rogers, T., and B. Warmke. 2015. Situationism versus situationism. Ethical Theory and Moral Practice 18(1): 9–26.

Rosenbaum J. E. 2006. Reborn a virgin: adolescents' retracting of virginity pledges and sexual histories. American Journal of Public Health 96(6): 1098–1103.

Ross, L., and R. E. Nisbett. 1991. The Person and the Situation: Perspectives of Social Psychology. New York: McGraw-Hill.

Rudy-Hiller, F. 2020. Reasonable expectations, moral responsibility, and empirical data. Philosophical Studies 177: 2945-2968.

Rushton, J. P., and A. C. Campbell. 1977. Modeling, vicarious reinforcement and extraversion on blood donating in adults: Immediate and long-term effects. European Journal of Social Psychology 7(3): 297–306.

Russell, D. 2009. Practical Intelligence and the Virtues. Oxford: Oxford University Press.

Russell, D. C. 2015a. Aristotle on cultivating virtue. In Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology, ed. N. E. Snow. Oxford: Oxford University Press. pp. 17-48.

Russell, D. C. 2015b. From personality to character to virtue. In Current Controversies in Virtue Theory, ed. M. Alfano. New York: Routledge. pp. 92-105.

Rust, J., and E. Schwitzgebel. 2013. Ethicists' and nonethicists' responsiveness to student e-mails: relationships among expressed normative attitude, self-described behavior, and empirically observed behavior. Metaphilosophy 44(3): 350–71.

Sabini, J., and M. Silver. 2005. Lack of character? Situationism critiqued. Ethics 115(3): 535–62.

Sarkissian, H. 2010. Minor tweaks, major payoffs: the problems and promise of situationism in moral philosophy. Philosophers' Imprint 10(9): 1-15.

Sartorio, C. 2018. Situations and responsiveness to reasons. Noûs 52(4): 796–807.

Scanlon, T. 2008. Moral Dimensions: Permissibility, Meaning, and Blame. Cambridge, MA: Belknap Press.

Schoeman, F. 1990. Psychology and standards of reasonable expectation. Public Affairs Quarterly 4(4): 387–402.

Schwitzgebel, E. 2009. Do ethicists steal more books? Philosophical Psychology 22(6): 711–725.

Schwitzgebel, E. 2013. Are ethicists any more likely to pay their registration fees at professional meetings? Economics and Philosophy 29(3): 371–80.

Shahade, J. 2015. On chess: physical fitness becomes increasingly important for top-level players. St. Louis Public Radio, 21 Jan. http://news.stlpublicradio.org/post/chess-physical-fitness-becomes-increasingly-important-top-level-players

Shoemaker, D. 2015. Responsibility from the Margins. New York: Oxford University Press.

Simon, H. A., and W. G. Chase. 1973. Skill in chess. American Scientist 61(4): 394–403.

Slingerland, E. 2011. The situationist critique and early Confucian virtue ethics. Ethics 121(2): 390–419.

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

Solomon, R. C. 2003. Victims of circumstance? A defense of virtue ethics in business. Business Ethics Quarterly 13(1): 43–62.

Sosa, E. 2009. Situations against virtues: the situationist attack on virtue theory. In Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice, ed. C. Mantzavinos. New York: Cambridge University Press. pp. 274-290.

Sreenivasan, G. 2002. Errors about errors: virtue theory and trait attribution. Mind 111: 47–68.

Sreenivasan, G. 2008. Consistency and character: still more errors. Mind 117(467): 603–12.

Stich, S. P. 1990. The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation. Cambridge, MA: MIT Press.

Stichter, M. 2007. Ethical expertise: the skill model of virtue. Ethical Theory and Moral Practice 10(2): 183–94.

Stichter, M. 2011. Virtues, skills, and right action. Ethical Theory and Moral Practice 14(1): 73–86.

Stichter, M. 2018. The Skillfulness of Virtue: Improving Our Moral and Epistemic Lives. Cambridge: Cambridge University Press.

Strawson, G. 1994. The impossibility of moral responsibility. Philosophical Studies 75: 5–24.

Swanton, C. 2003. Virtue Ethics: A Pluralistic View. Oxford: Oxford University Press.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. Science 211: 453–8.

Upton, C. L. 2009. The structure of character. Journal of Ethics 13(2–3): 175–93.

Upton, C. L. 2017. Meditation and the cultivation of virtue. Philosophical Psychology 30(4): 373–94.

van Der Maas, H. L., and E. J. Wagenmakers. 2005. A psychometric analysis of chess expertise. American Journal of Psychology 118(1): 29–60.

Vargas, M. 2013a. Building Better Beings: A Theory of Moral Responsibility. Oxford: Oxford University Press.

Vargas, M. 2013b. Situationism and moral responsibility: free will in fragments. In Decomposing the Will, ed. T. Vierkant, J. Kiverstein, and A. Clark. New York: Oxford University Press. pp. 325-349.

Vargas, M. 2017. Implicit bias, moral responsibility, and moral ecology. In Oxford Studies in Agency and Responsibility, vol. 4, ed. D. Shoemaker. New York: Oxford University Press. pp. 219-247.

Vargas, M. 2018. Reflectivism, skepticism, and values. Social Theory and Practice 44(2): 255–66.

Vargas, M. 2020. Negligence and social self-governance. In Surrounding Self-Control, ed. A. R. Mele. New York: Oxford University Press. pp. 400-420.

Vargas, M. Forthcoming. Instrumentalist theories of moral responsibility. In The Oxford Handbook of Moral Responsibility, ed. D. Nelkin and D. Pereboom. Oxford: Oxford University Press.

Verhoeven, A. A. C., M. A. Adriaanse, D. T. D. de Ridder, E. de Vet, and B. M. Fennis. 2013. Less is more: the effect of multiple implementation intentions targeting unhealthy snacking habits. European Journal of Social Psychology 43: 344–54.

Vinkers, C. D. W., M. A. Adriaanse, F. M. Kroese, and D. T. de Ridder. 2015. Better sorry than safe: making a Plan B reduces effectiveness of implementation intentions in healthy eating goals. Psychology and Health 30: 821–38.

Vranas: 2005. The indeterminacy paradox: character evaluations and human psychology. Noûs 39(1): 1–42.

Waggoner, M. 2021. The focus of virtue: Broadening attention in empirically informed accounts of virtue cultivation. Philosophical Psychology 34(8): 1217-1245.

Washington, N., and Kelly, D. 2016. Who's responsible for this? Moral responsibility, externalism, and knowledge about implicit bias. In. Implicit Bias and Philosophy, vol. 2, ed. J. Saul and M. Brownstein. Oxford: Oxford University Press. pp. 11-36.

Watson, G. 1975. Free agency. Journal of Philosophy 72(8): 205–20.

Webber, J. 2006. Character, consistency, and classification. Mind 115(459): 651–8.

Wegner, D. M. 2002. The Illusion of Conscious Will. Cambridge, MA: MIT Press.

Wielenberg, K. 2006. Saving character. Ethical Theory and Moral Practice 9: 461–91.

Wolf, S. 1990. Freedom Within Reason. New York: Oxford University Press.

Zagzebski, L. 2017. Exemplarist Moral Theory. Oxford: Oxford University Press.