

ARISTOTELEAN VIRTUE AND THE INTERPERSONAL ASPECT OF ETHICAL CHARACTER

ABSTRACT

I examine the Aristotelean conception of virtuous character as firm and unchangeable, a normative ideal endorsed in the currently influential, broadly Aristotelean school of thought known as 'virtue ethics.' Drawing on central concepts of Aristotle's *Nicomachean Ethics*, I offer an account of how this ideal is supposed to be realized psychologically. I then consider present-day empirical findings about relevant psychological processes, with special attention to interpersonal processes. The empirical evidence suggests that over time, the same interpersonal processes that sometimes help to sustain character may also disrupt it, even among agents who have the right values in principle. Fortunately, the evidence also suggests some remedial measures. An important philosophical measure, I conclude, is for advocates of virtue ethics to address agents' psychological need for a systematic decision procedure that will focus attention primarily on substantive ethical considerations, rather than characterological assessment.

KEYWORDS

Aristotle, character, moral psychology, situationism, virtue, virtue ethics

ARISTOTELEAN VIRTUE AND
THE INTERPERSONAL ASPECT OF ETHICAL CHARACTER¹

Maria Merritt

Aristotle holds that an agent acts virtuously only if his choices of virtuous action “proceed from a firm and unchangeable character.”² That is, the agent should be firmly established in his dispositions to respond emotionally in the right way to the right objects of choice and avoidance, to deliberate well, and to choose the right actions for the right reasons. Aristotle’s normative ideal of firm and unchangeable character is endorsed in the currently influential, broadly Aristotelean school of thought known as “virtue ethics.”³

Drawing on central concepts of Aristotle’s *Nicomachean Ethics*, I offer an account of how the normative ideal of firm and unchangeable character is supposed to be realized psychologically (Section II). I then consider this aspect of Aristotelean virtue in light of present-day empirical findings about relevant psychological processes (Section III). Since the Aristotelean conception of virtue emphasizes consistency and integration within the agent’s set of motivations, an empirical research topic of great interest is the psychological processes by which individuals regulate themselves – or fail to regulate themselves – with reference to their normative commitments.

The evidence suggests that these processes have an interpersonal aspect. Typically, and often not consciously, their operation will be mediated by our social bonds with other individuals whose values we take ourselves to share. This proposition might seem to sit well with a reasonably nuanced account of Aristotelean virtue. Aristotle allows that among those who share the end of living a good life, social bonds like

friendship can help individuals to sustain virtuous character. But trouble emerges in the form of further empirical findings. The same interpersonal processes that sometimes help to sustain character may also disrupt it, even among agents who have the right values in principle. The character-oriented self-understanding that would-be virtuous agents share with their fellows can give rise to misplaced confidence, on the part of each individual, that his or her normative commitments will reliably issue in right action, even while an objective viewpoint reveals that sometimes they do not. Moreover, ironically, the more importance we place on seeing ourselves as aiming at the good, the more vulnerable we may be to this insidious psychological dynamic.

To put it bluntly, the problem is not whether you have the right values but whether you act in accordance with the values you have, and how interpersonal influences that are otherwise ethically constructive may, over time, pull your actions apart from your professed values. Fortunately, the empirical evidence also suggests some remedial measures, both practical and philosophical, to counteract the potentially damaging effects of these influences. The practical measures include a stance of due fallibilism and modesty about the state of one's own character, caution against over-idealizing the character of close associates, and a commitment to building mechanisms of objective accountability into institutions and organizations. An important philosophical measure, I conclude (Section IV), is for the advocates of virtue ethics to address agents' psychological need for a systematic decision procedure that focuses attention primarily on substantive ethical considerations rather than characterological assessment. In order to put virtue ethics effectively into practice, agents need a reliable way to track and organize the many substantive ethical considerations that are theoretically unified in the abstract

conception of ‘what the virtuous person would do.’

I. Moral Theory and Scientific Psychology

We need first to address a methodological concern: why and in what ways does experimental scientific psychology matter for philosophical moral psychology?

Philosophers in growing numbers have taken up the problem of compatibility between the two.⁴ This is not merely a matter of general-purpose intellectual responsibility, an interest in keeping one’s home discipline informed by relevant results from other disciplines. The reason why moral philosophers should aim for compatibility with scientific psychology is specific to the purposes of normative theory. In general, any normative theory has to take account of the empirical understanding – if there is one – of what it applies to. Where X is a kind of object (or state of affairs or event) in the natural world, a theory intended as normative for the behaviour (or occurrence) of X calls for an empirically informed account of what X is.⁵

For example, since moral theories purport to articulate norms for human action, they should be informed by an empirical understanding of what actions are, and of what people are doing when they engage in behaviour that folk psychology takes to be intentional. Similarly, as I aim to illustrate in this article, moral theories that purport to articulate a psychology of virtue – a set of norms for morally important dispositions of personal character – should be informed by an empirical understanding of the psychological processes that affect the continuity and sustainability of such dispositions over time.

Those who resist the intrusion of scientific psychology into moral theory might

couch their opposition as follows. First, since the primary method of scientific psychology is controlled experimentation, whose results are at best inductive generalizations, it cannot establish the impossibility of exceptions. Second, it is perfectly in keeping with the normative purposes of moral theory to exhort us towards an ideal that is only ever realized in exceptional cases, cases of precisely the kind that inductive generalizations leave room for.

This response contains a kernel of methodological truth. Since the purposes of moral theory are primarily normative, not predictive or explanatory, compatibility with scientific psychology is at most a constraint on moral theory, not a goal in its own right. Where to invoke the constraint depends in part upon what empirical commitments a given instance of moral theory incurs as it goes about its normative business. A moral theory might tailor its psychology of virtue to normative purposes by articulating an ideal of practical rationality and knowledge of the good, while recognizing that the ideal is rarely attained, and may well be unattainable by the vast majority of men and women. Such a moral theory might then incur no commitment to constrain its normative psychology of virtue by the inductive generalizations of scientific psychology, although it could use them to diagnose and inventory the ways in which people are liable to fall short of the ideal.

Rachana Kamtekar has argued that Plato, Aristotle, and the Stoics take something like this approach to the psychology of virtue.⁶ By the same reasoning, Kamtekar and other advocates of traditional virtue ethics can claim that their approach is compatible with some of the seemingly problematic findings of experimental psychology. The findings that have motivated virtue ethicists to rally most vigorously to this defence are

those of ‘situationist’ social psychology, which suggest that most people’s morally important behaviour is often better explained by chance situational factors than by dispositions of personal character. Even John Doris, who has presented at length the situationist argument against routine attribution of virtuous (or any) character traits, acknowledges that situationism leaves open the possibility of individual virtue in rare cases.⁷ Eric Hutton, in sympathy with traditional virtue ethics, has recently interpreted early Confucian thought as compatible with situationist social psychology in just this way. While the Confucians recognize that ordinary people are susceptible to the vagaries of situational influence, at the same time they believe in rare “gentlemen and sages whose good behaviour is utterly stable and consistent.”⁸

Yet the social function of moral norms is to regulate actual conduct among ordinary men and women in human populations. And part of the point of philosophical moral theory is to reflect critically on what the content of moral norms should be, given their social function. If traditional philosophical ideals of virtue are so lofty that even their advocates expect them to be realized only rarely at best, how can they be relevant to practical morality?

One possibility is that individuals might strive toward such ideals as aspirational goals.⁹ Even when the goal is unattainable you can hope, through the effort of trying to reach it, to become better than you might otherwise have been. A second possibility is to adopt a division-of-labour model and hold out hope for a moral meritocracy to keep the rest of us in line. Hutton suggests that in early Confucian thought, reflection on ideals of virtue could be crucial not despite the rarity of virtue, but because of it. Given a traditional hierarchical society, it is ideally the virtuous “gentlemen” and “sages” who

will establish the social structures that ordinary (*ex hypothesi*, mostly non-virtuous) people inhabit: “if there are people who do have robust character traits and are resistant to situational variation, they can design and reliably maintain the broad range of institutions and situations that facilitate good behaviour for everyone else.”¹⁰ Transposing the idea to modern liberal democracies, we might hope that the individuals who lead particular communities, organizations, and institutions will turn out to be the ones who prove capable of virtue.¹¹

These two possibilities are not mutually exclusive. Individuals could be encouraged to aspire to virtue whether or not they hold positions of social or political leadership. Those who do hold such positions could be given to understand that their institutional roles carry an extra burden of responsibility to strive for virtue, commensurate with the importance of their influence on situational settings that guide the behaviour of men and women subject to their authority. But even if a few virtuous super-individuals exist, there is no particular reason to be confident that they will regularly occupy positions of power (not least because in order to secure such positions they would have to compete against individuals and organizations who value power above all). In general, if virtue is so rare as traditional virtue ethicists are willing to maintain, and if our behaviour is influenced by morally arbitrary factors so extensively as they are willing to admit, we have good reason to doubt that we should entrust the effective social functioning of moral norms to the personal resolve and conscientious striving of individuals (ourselves or others, highly placed or not).

My concern in this article is with the psychology of interpersonal processes as they affect morally important personal dispositions over time. This issue is distinct from

the problem highlighted by situationism, which is that morally irrelevant factors affect morally important behaviour at the time of action. Focusing on the influential Aristotelean ideal of virtuous character as firm and unchangeable, I argue that the evidence about interpersonal processes and their effect on personal dispositions over time provides a distinct and additional reason, over and above situationism, to predict that ordinary men and women will run into trouble in the attempt to realize such an ideal.¹² The safest bet is to assume that however well-intentioned and conscientious we may be, empirical generalizations about interpersonal processes apply to each of us – that is, to each of ‘us’ who take an interest in morality when we make choices about how to act, how to live our lives, and how to shape the social settings for which we bear some responsibility.

II. Firm and Unchangeable Character: Moral Psychology and Development

How, in Aristotle’s ethics, is the normative ideal of firm and unchangeable character supposed to be psychologically realized? What is essential to Aristotelean virtue is the agent’s choice of some actions and rejection of other actions, under descriptions that are sensitive to relevant ethical considerations.¹³ This is one of the points at which the current orthodox philosophical conception of virtue is directly indebted to Aristotle, who defines it as a *hexis prohairetike*, a “state of character (*hexis*) concerned with choice (*prohairetike*).”¹⁴ In Aristotle’s conception a vice or a virtue, as a *hexis prohairetike*, is a disposition to respond in particular situations by choosing certain actions. Virtue is the disposition to respond by choosing the right actions.¹⁵ Virtue can be analyzed into two basic components, practical wisdom (*phronesis*) and moral virtue

(*ethike arete*). As Aristotle conceives of them, each component tends to perpetuate itself, and each tends to reinforce the other as their activity continues over the course of the agent's life. Once acquired, the motivational structure of virtue thus tends to perpetuate itself, making virtuous character firm and unchangeable. As Aristotle puts it, the virtues "tend...by their own nature, to the doing of the acts by which they are produced."¹⁶

A. Practical wisdom

Practical wisdom is the ability to excel in deliberation about what is truly valuable, from the standpoint of reflection on the shape of your life as a whole. It involves an astute grasp, well-informed by experience, of how such goods as friendship, honour, achievement, wealth, justice, physical security, sensual gratification, intellectual activity, and so forth, are in play under various kinds of circumstances. An agent who has practical wisdom knows how to deliberate well and make good choices of action where these goods are concerned, in all situations.

Aristotle holds that practical wisdom cannot be forgotten.¹⁷ Once you have it you are not going to lose it, short of the most dire debilitation or destruction of your cognitive faculties. Suppose it is true that certain goods in life are the best goods it is possible for human beings to enjoy. Then once you have learned how to recognize them even in their most subtle appearances, and how to deliberate so that through your choices of action you do all you can to achieve them, not only in familiar situations but in perplexing or novel situations as well, that does not seem to be the kind of thing at which you can fall out of practice.

The kinds of choices practical wisdom helps us to deliberate about, choices about

how to act with respect to the goods of life, are always and everywhere with us. We have to make choices about these goods constantly, although it may be possible to get out of practice at making choices about some specific kind of good or other. For instance, if you spent several years in seclusion engaged in excellent intellectual activity, you might get out of practice at the kind of high-pressure decision-making needed for excellent political activity. But even without constant practice at deliberating about all the specific kinds of human goods, you would not lose touch with more fundamental practical insights, say, about the comparative value of various goods. You always have the opportunity to practice using such insights, given that you always have choices to make with respect to at least some of the goods of life. At least, this is so for as long as you continue to possess moral virtue.

B. Moral virtue

Moral virtue is excellence in the activity of the emotions and appetites. It keeps you in the attitude of caring appropriately about the goods of life, so that your concern for getting your choices right about those goods will guarantee you a succession of appropriate occasions for the exercise of practical wisdom. In Aristotle's view the life of virtue really is the best life, and so long as the appetitive domain of your psyche is in a condition to experience it as best, you can be counted on to stick with that life.

More precisely, excellence in the appetitive domain is foundational to Aristotle's characterization of the virtuous man as one who is such as to act for the sake of the noble. Aristotle claims that there are in general three objects of choice and three objects of avoidance: "the noble, the advantageous, the pleasant, and their contraries, the base, the

injurious, the painful.”¹⁸ At the same time, pleasure “accompanies all objects of choice; for even the noble and the advantageous appear pleasant.”¹⁹ Some objects of choice are merely pleasant: they appeal to any creature subject to sensual appetites. But other objects of choice are pleasant because and only so far as the agent is able to relate them to goods that take a certain kind of rational awareness to appreciate, goods which are for that reason superior. Rest following extreme physical exertion is merely pleasant; running hard for miles to get an urgent message to the commanding officer in a battle is pleasant, or rather is rightly to be found pleasant, because the honour of serving well in battle is noble. With respect to the objects of avoidance, running hard for miles is merely painful, whereas it would be disgraceful to arrive too late with the message because, feeling tired, one slowed down to a comfortable jog. As a disgraceful action, slowing down should strike the virtuous man as far more repellent – and repellent in a more important way -- than continuing to run as hard as he can over the necessary distance, physical pain notwithstanding. In general, the virtuous man has accurate sensibilities and the right priorities when it comes to such distinctions among objects of choice and avoidance. He recognizes what is noble and values it above all, and he recognizes what is base and reviles it more strongly than he does anything else.

Moral virtue, then, is constituted by our values in the sense of our attractions and aversions to the appropriate objects of choice, existing in harmony with our appetites, emotions, and actions. Once the development of moral virtue initially establishes the attitudes of loving the noble and despising the base, we will energetically seek to learn more about how to make practical choices in accordance with those attitudes, progressing in the gradual acquisition of practical wisdom. Abilities that we develop on the way to

moral virtue, such as impulse control and other forms of self-control necessary for keeping emotional and appetitive activity in line with our values, are in due course supposed to reach beyond mere self-control to culminate in the highly integrated self-mastery of virtuous character.

Meanwhile the process of acquiring practical wisdom teaches us more about the nature of the nobler objects of choice, thereby adding depth, intensity, and discerning precision to our love for the noble. What constitutes practical wisdom is our commitment to our values in reflective understanding, manifested in our intelligent recognition of when and how they are in play on occasions of deliberation and choice, including circumstances of adversity. The continued accumulation of experience gradually reveals more and more of the ways in which the nobler objects of choice really are superior, keeping us in the attitude of caring about getting our choices right.

C. The interpersonal aspect of virtuous character

Aristotle allows that even mature virtuous character needs at least some support from social relationships. We can begin by remembering that in his view the whole value of virtue lies in its active exercise, as distinct from its mere possession.²⁰ Now add the consideration that a great many of the Aristotelean virtues -- courage in battle, justice, liberality, magnificence, pride, due ambition, good temper, friendliness, truthfulness, ready wit -- require for their active exercise a social world inhabited by a community of peers. This is so for reasons over and above Aristotle's argument that since human beings are social animals, we can live the best human life only in society. More interestingly for our purposes, it is so because Aristotle's conception of virtue, founded in love for the

noble and contempt for the base, is highly sensitive to standards of honour.

For Aristotle as for Homer, whom he quotes approvingly on this point in his discussion of courage, honour is an essentially social phenomenon.²¹ In order for a man to be fully assured that in his conduct he meaningfully satisfies a standard of honour, he needs to win the recognition and respect of other men who are honourable themselves. And for their part, each of them needs the same.²² The standard of honour must be located in the judgment of fellow men of practical wisdom, as opposed to the judgment of anyone at all, however wise or foolish. The value of honour is in its guidance and confirmation of the honoured man's understanding of himself as virtuous; only honour bestowed by men of practical wisdom can deliver this result.²³

Similarly, in Aristotle's account of friendship, friends are supposed to love each other on the basis of virtue, each loving the other for his good character. They choose to live together and they coordinate their activities to pursue various specific ends that they share. The overall end that unites them is the end of living a good life. One reason they can help each other to pursue this larger end is that each admires and tries to emulate the other. Given that I regard my friend as someone generally of good character, a person practiced in living well and better at it in some respects than I am, I can learn from my friend how to improve in those respects. And in the best kind of friendship there will be reciprocity: my friend can learn from me in the same way, *mutatis mutandis*.²⁴

So, in Aristotle's view, the virtuous man needs a social world for the expression and exercise of the virtues. He needs not just any social world, but one that prominently features social bonds with others who share his evaluative sensibilities and priorities.

III. Empirical Psychology: Interpersonal Processes and the Discontinuity of Character

It is a commonplace of Aristotelean moral psychology that in early life the agent will, ideally, begin to adopt the values characteristic of virtue by way of training and education in a supportive social community. With increasing maturity and reflective understanding comes increasing independence in the agent's commitment to the right values. The goal of living up to the right values, initially a goal assigned by others, is transformed into a goal we set for ourselves. But even self-set goals may continue to be influenced by the continuing normative expectations of selected others: those who share our values, and with whom we choose to maintain interpersonal relationships. As noted above, this is what we find in Aristotle's ideal of the virtuous man's social world.

Using the resources of present-day scientific psychology, I shall suggest how interpersonal processes operate through social bonds with like-minded others, with the effect of helping individuals to sustain virtuous normative commitments over time (Section A); I shall then give an empirically informed account of how certain aspects of the same interpersonal processes also threaten to break the self-reinforcing cycle between moral virtue and practical wisdom, prompting occasions of discontinuity that may over time become lasting disruptions (Section B). Rather than simply serving to strengthen character, as the Aristotelean view would have it, social bonds with others who share our evaluative sensibilities and priorities – even assuming we have got those right in principle – may also tend to undermine character.

A. Self-Evaluation and Self-Presentation

The agent's values are the psychological ground of virtue. For Aristotelean virtue, we may more accurately speak of positive and negative values, or values and anti-values, since virtue requires getting it right not only about the noble as an object of choice, but also about the base as an object of avoidance.²⁵ This motivational structure brings with it a self-evaluative concern to be, or to become, a certain kind of person.

In the first place, the very activity or stance of valuing draws some of one's attention to self-evaluation. As Agnieszka Jaworska has pointed out, "A person values *herself* in terms of how well she lives up to her values."²⁶ On top of that, the standard Aristotelean tradition specifically encourages the use of broadly characterological attributions in evaluating one's own and others' actions. Indeed, for Aristotle the normative criterion of virtuous choice, the *orthos logos*, is set by the idea of choosing as the man of practical wisdom would choose.²⁷ Similarly, contemporary neo-Aristoteleans like Hursthouse set out to define right action in terms of "what a virtuous agent would characteristically (i.e. acting in character) do in the circumstances."²⁸ By implication, characterological self-evaluation is central to the first-person practice of virtue ethics in the Aristotelean tradition.

Psychologists have studied a similar activity under the name of 'self-discrepancy theory.' This is a topic under the broader heading of 'self-regulation,' the study of behaviour as a purposive, continually self-adjusting process of moving toward desired outcomes and away from undesired outcomes.²⁹ According to self-discrepancy theory, people regulate their behaviour in part by reference to 'self-guides,' to which they compare their perceptions of their actual selves in an ongoing effort to become more like a 'desired self.'³⁰

Interpersonal processes that engage with self-regulation to influence subsequent behaviour constitute a major research topic in social psychology.³¹ Experiments reveal the cognitive, affective, and behavioural manifestations of individuals' interest in living up to the normative expectations of selected others, whose responses can serve as benchmarks and provide external feedback to help guide the process of becoming more like one's desired self.³² A pervasive type of behavioural manifestation is 'self-presentation' (also known as 'impression management'), whereby each participant in a social encounter shapes the impression he or she makes on others. Far from necessarily involving manipulation or any sort of conscious calculation, self-presentation may often be neither accessible to conscious awareness nor controlled by conscious effort. And far from being confined to the likes of shady used-car salesmen, it is basic to successful communication and social interaction, not only for first encounters but also over the long run in relationships of collegiality, friendship, or love. Its interpersonal function is to allow us to select and streamline the information we convey about ourselves, in a manner tailored to the relationship and circumstances at hand.³³

In many interpersonal contexts, self-presentation can have the effect of encouraging or confirming others' beliefs that we possess dispositions of character that they (and we) consider desirable.³⁴ Over time, when we engage in repeated interactions with the same parties they acquire expectations about how we will act not only during one encounter, but also on future occasions. If we care about maintaining our relationships, we should generally prefer to satisfy such expectations (or to renegotiate them so that it's easier to satisfy them in the future).³⁵ It makes sense to act accordingly, at least when it is reasonable to assume that others stand in a basically cooperative

attitude toward us. The cumulative effect can be to establish and sustain dispositions of motivation and choice that tend to produce the desired pattern of conduct.

For instance, studies of ‘the Michelangelo phenomenon’ have found that couples in close-partner relationships “sculpt” each other’s “dispositions, values, and behavioural tendencies,” through the expectations that each forms of the other. In a stable, well-functioning relationship, each partner endorses the other’s desired self, and communicates expectations that tend to bring out the best in the other with reference to this ideal. (In unhappy relationships the process backfires, each partner pushing the other further away from his or her desired self.)³⁶ Similarly, as we maintain interpersonal relationships and fulfil our social roles over an extended period of time, the cumulative effect may be to reinforce in each of us the personal dispositions that are most closely bound up with the expectations we have a stake in fulfilling.

With regard to ethical practice, I suggest that whatever success we may enjoy in meeting our own self-evaluative standards for personal character, even in maturity and even assuming we have somehow really got the right values, is typically subject to our ongoing engagement with the expectations of particular other individuals with whom we hold our values in common. What I mean to highlight here is nothing as superficial as an immature, unreflective dependence upon others for overt prompting about how to act, or a lack of motivation to act rightly unless those others are currently present and looking on.³⁷ The reinforced dispositions can issue in motivation and choice straightforwardly, and no less so in situations of which one believes not another living soul will ever learn what happened. My point is that typically, these deep springs of action may themselves be sensitive to ongoing reinforcement from interpersonal processes.

B. Moral Misapprehension and Differential Accessibility

This picture has darker shadows. Empirical evidence suggests that the very interpersonal processes which ideally support ethical character also expose it to discontinuity and disruption, in ways that may be extremely difficult for the agent to detect. To illustrate, consider a hypothetical example featuring justice.

Angela works in the public defender's office of a large city, with colleagues she respects and likes. She went to work there in the first place, and stayed, because she saw that her colleagues shared her values. In her conduct, Angela consistently demonstrates a strong commitment to the ethical principle that justice requires equal treatment under the law. Where justice for indigent defendants is concerned, she dedicates herself to making sure that the criminal justice system delivers to them the treatment to which they are entitled by law. She does so in the right spirit (caring that justice be served), and for the right reason (that justice requires equal treatment). In short, Angela's values, choices, and reasons for action are unimpeachable with respect to justice.

But over time, the group's caseload gradually increases. In their efforts to keep up, they begin to cut corners, but in each case they manage to rationalize their actions as ethically unproblematic. Without quite realizing it, Angela too makes more and more choices that are in fact inconsistent with what justice requires, going in for her share of well-meaning but unwittingly self-serving rationalizations. Angela continues to understand herself as dedicated to justice, but in her actual conduct she is no longer just. Indigent defendants suffer the consequences.

If there is a single way to characterize the downward spiral of mistakes that Angela and her colleagues have fallen into, it is misapprehension of the objects of choice

and avoidance. In terms of Aristotle's account of acting for the sake of the noble with respect to the objects of choice and the objects of avoidance, they have become prone to choose actions they ought to avoid – actions that are advantageous but wrong (base) – under misapprehensions of them as advantageous and right (noble).

A sadly compelling real-life example is the scandal that recently surrounded Patricia Dunn, former Chairman of the Board at the Hewlett-Packard Company (HP). Dunn is an expert in corporate governance. When she joined the HP board, she brought with her an impeccable record of ethical responsibility in powerful positions such as sole global C.E.O. of Wells Fargo Investment Advisors, where she had been “the principal fiduciary for more than a trillion dollars in assets.”³⁸ As board chairman at HP, in a well-intentioned campaign to contain high-level leaks to the press, Dunn trusted her in-house legal and ethics team to determine the lengths to which hired private detectives could go in their investigations of board members.³⁹ Unfortunately, the detectives went so far as to draw felony counts of fraud and identity theft, not only against themselves but also against Dunn.⁴⁰ At issue was the practice of “pretexting”: assuming a false identity to get confidential information about targets of investigation. On a plausible interpretation, Dunn and her closest confidantes were so preoccupied with their goal of containing leaks – itself an entirely above-board objective, which Dunn was specifically hired to pursue – that they all too readily overlooked the base and shameful (not to say criminal) aspects of the actions taken under their authority.

Research in social psychology strongly suggests that such misapprehensions are an ordinary and predictable result of human cognitive functioning. They are ‘ordinary’ in two senses: first, they are generated by the routine operation of normal cognitive

processes; second, far from being peculiar to a few unusual individuals whose character is rotten through and through, ethical failures of this kind are ubiquitous, even among well-meaning, intelligent, sensitive people.⁴¹ Be our ethical commitments ever so correct and thoughtfully adopted, we are (virtually) all susceptible to choosing seemingly (to us) innocuous actions which objective observers would easily recognize as instances of ethical failure, such as conflict of interest, in-group favouritism, or stereotyping and prejudice.⁴² For example, several findings about physicians' prescribing practices are well-supported by empirical research: first, in aggregate, physicians' prescription practices are significantly influenced by even minor pharmaceutical promotions (such as providing pens and notepads emblazoned with the corporate logo); second, a majority of physicians self-report that their practices are not influenced by such promotions; and third, an even greater majority believe that other physicians, unlike themselves, are indeed influenced by promotions.⁴³ Together, these findings suggest that individuals can easily recognize conflict of interest in their peers' behaviour while being oblivious to it in their own.

The explanation for such obliviousness is that many less-than-ideal choices of action issue from psychological processes that operate largely outside conscious awareness, selectively determining what comes most readily to mind and so directing behaviour in ways that may run counter to the subject's reflectively endorsed values and ethical self-understanding.⁴⁴ Part of the empirical and theoretical basis for this explanation is found in the longstanding research program on intuition, judgment, and decision-making led by Daniel Kahneman and the late Amos Tversky, now carried on by Kahneman.⁴⁵ Kahneman's explanatory framework emphasizes differential 'accessibility'

– “the ease (or effort) with which particular mental contents come to mind” – among various mental contents relevant to the judgment or decision at hand.⁴⁶ In the following account, I will first discuss patterns of differential accessibility and their consequences for ethically important behaviour. Then I will turn to the causal contributions made by several factors, especially interpersonal processes, which operate largely outside the purview of conscious awareness to determine different degrees of accessibility.

Consider the misapprehension of an advantageous action as morally right when it is actually wrong. The mental contents that come easily to mind track the features of the situation that recommend a certain action as advantageous. The mental contents that do not come to mind would (if they did come to mind) track the features of the situation that might make that action wrong. These missing mental contents would have prompted the agent to seek further information, entertain alternative courses of action, weigh the trade-offs that might arise between important considerations, and generally to engage in critical, reflective deliberation. But since they don't come to mind, the agent fails to consider them, so that her choice of action is uninformed by them. The action then appears to be both advantageous and, at worst, morally neutral. So, the agent chooses the advantageous action under an erroneous description of it as morally acceptable.⁴⁷

The initial error of choosing an advantageous action under a misapprehension of it as morally acceptable can be seriously compounded by overconfident faith in one's own good intentions.⁴⁸ Broad characterological self-attributions further increase the hazard. The belief 'I am a good person,' or even 'I am trying my best to be a good person,' encourages the expectation 'I would not deliberately do anything wrong.' Even if this expectation were accurate, its content is quite distinct from that of a further expectation

that can too easily slip into its place during the everyday flow of the first-person perspective on our experience and conduct: ‘I would not do anything wrong, deliberately or not.’ Even though we might not openly, in full awareness, credit ourselves with moral infallibility – that would be immodest – we are prone to limit our field of vision unwittingly, so that the effective reach of critical self-monitoring extends only to those occasions when we do take time to think carefully. Left out of consideration are occasions when everything seems fine, and we are not in fact doing wrong deliberately, but had we deliberated we would have noticed decisively wrong-making features of our course of action. Instead, if and when we consider the latter type of occasion in retrospect, the conviction of being a good person tends to displace deliberative effort into after-the-fact rationalization, a selectively confirmatory survey of all the reasons that seem to support the past decision. In hindsight, we may even enhance our description of the action as not only morally neutral, but downright laudable.⁴⁹

Through such unfortunate trains of cognitive processes, we can inadvertently acquire a new predisposition to continue misapprehending future situations and actions of a similar type. In Aristotelean terms, this is the piecemeal corrosion of practical wisdom, or, hardly less worrisome, the piecemeal corrosion of its building blocks (in the case of agents who are only partly in possession of incipient practical wisdom). The practically wise agent knows how to estimate the comparative value of all relevant goods and evils as they come into play in any particular situation. It is part of practical wisdom, and a necessary step on the way to acquiring it, that we should learn to estimate the value of specific goods and evils as they crop up, one by one or cluster by cluster, in the local practical milieu we happen to inhabit. When, instead, faulty estimations of some specific

goods and evils become habitual, we may continue to profess a generalized love of noble actions and revulsion toward base actions, yet inadvertently decouple those attitudes from some of their correct objects of choice and avoidance. In this way, our aspirations toward moral virtue can pull apart from true practical wisdom. The cycle of self-reinforcement between moral virtue and practical wisdom is broken.

Moreover, the overconfident conviction of one's own moral probity can be bolstered by interpersonal dynamics – which may be otherwise ethically salutary, as we saw in part (A) above – that shape, encourage, and support one's self-understanding as a person of good character. Experimental studies have found that subjects' accounts of their actions are more likely to be distorted by self-enhancing cognitive bias when their actions are known to others and they feel responsible for the outcome.⁵⁰ In the words of one pair of psychologists, “The presented self is (usually) too good to be true; the (too) good self is often genuinely believed.”⁵¹

Not only do such interpersonal factors exert an unhelpful influence downstream of the faulty judgments that agents initially make under conditions of differential accessibility. More importantly, they contribute directly to differential accessibility in the first place, by influencing what kind of cognitive processes are likely to direct an agent's behaviour.

Different degrees of accessibility between mental contents are determined jointly by two types of causal factors: the characteristics of the cognitive processes that produce those mental contents in one's mind, and the “characteristics of the stimuli and events” that one encounters in the situation at hand.⁵² Let us look first at cognitive processes. In the kinds of ethical failure that psychologists have studied most extensively (such as

conflict of interest and in-group favouritism), the cognitive processes that bring the action's advantageous features to mind are "typically fast, automatic, effortless, associative, implicit (not available to introspection), and often emotionally charged."⁵³ The cognitive processes that would bring the wrong-making features to mind (but do not, because they do not occur at the right moment) are "slower, serial, effortful, more likely to be consciously monitored and deliberately controlled."⁵⁴

What is missing in many cases of misapprehension is the engagement of the slower, deliberative type of process to examine the situation with some critical distance, search for further relevant information, reason through important considerations, and in general to monitor or check the choice of action prompted by the faster, effortless type of process. The capacity for sound ethical deliberation is of little use if we fail to employ it on occasions that call for it. As the psychologist John Darley puts it, "people are ethical, but only intermittently so."⁵⁵ Enter the second type of causal factor in the explanation of differential accessibility: characteristics of stimuli and events in the situation at hand. Given that the slower, deliberative type of cognitive process is often necessary to bring relevant moral considerations to awareness, what conditions tend to stultify it? And what conditions tend to awaken it?

Psychological studies of accountability in judgment and decision-making open a revealing window onto differential accessibility. 'Accountability' is defined, for the purposes of assessing these studies, as "the implicit or explicit expectation that one may be called upon to justify one's beliefs, feelings, and actions to others."⁵⁶ It turns out that accountability *simpliciter* may or may not awaken the effortful cognitive activity of deliberation. Whether it will do so depends, in part, on the expected characteristics of the

audience to whom one is accountable. Numerous experimental and field studies have shown that expecting to be accountable to an audience whose views one knows ahead of time will stultify cognitive effort, as compared with expecting to be accountable to an audience whose views one does not know.⁵⁷

When experimental subjects are asked to make judgments or decisions, and they know the preferences of the audience who will evaluate the result, their behaviour conforms to the use of cognitive shortcuts guided by their beliefs about what the audience will find acceptable, bypassing the cognitive effort required to assess the object of judgment independently.⁵⁸ Subjects who do not know the preferences of their audience, by contrast, anticipate the need to justify their judgment or decision, prompting a critical, reflective internal dialogue that awakens deliberation. These subjects more readily tolerate evaluative complexity and exercise care in considering complex information, weighing the reasons for and against alternative options, and recognizing trade-offs. Accountability to an audience with unknown views has been shown significantly to reduce subjects' tendency to proceed with undue haste and substitute oversimplified, superficial, or misleading decision procedures for effortful, systematic, self-critical deliberation. It has also been shown to attenuate the effects of a wide variety of cognitive biases, such as oversensitivity to the order in which information appears.⁵⁹

But in practice, we cannot avoid making many of our judgments and decisions under conditions in which we are thoroughly familiar with the views of others to whom we are accountable. We often hold important values, interests, and goals in common with them. Is there any way to keep ourselves cognitively awake and alert under these conditions? Further studies of accountability have found that the potentially stultifying

effect of having an audience with known views can be mitigated by inducing a motivational orientation toward accuracy in the decision-maker, as contrasted with a motivational orientation toward social compatibility.⁶⁰

For instance, one study manipulated the conditions under which subjects had to make a business decision between two options for allocating a \$10 million marketing budget: stay with a past decision that had yielded disappointing results, or initiate an alternative likely to improve the return on investment.⁶¹ Between two groups of subjects, one group induced to be motivated by accuracy and the other induced to be motivated by social compatibility, some subjects in each group were told that they would have to justify their recommendations in a face-to-face meeting with another person who endorsed staying with the past, disappointing decision. Other subjects were told about this other person's views, but were assured they would never have to meet the person, nor would they have to explain and justify the present decision. Thus there were four combinations of experimental conditions: accuracy motivation plus accountability to an audience with known views, compatibility motivation plus accountability to an audience with known views, accuracy motivation plus no accountability, and compatibility motivation plus no accountability.

The results were that in all but one of the experimental combinations, subjects tended to conform to the known views of the other person, to the detriment of return on investment. Only under conditions of accuracy motivation plus accountability did subjects escape the influence of the other person's known endorsement of a plan to throw good money after bad. These were the only subjects whose decision-making was comparable to that of control subjects who had no exposure whatsoever to anyone else's

recommendations. An explanation for such results in terms of differential accessibility is suggested by other studies showing that subjects with a social compatibility motivation are more likely to rest content with the low-effort cognitive shortcut of mirroring a discussion partner's opinions, while subjects with an accuracy motivation are more likely to engage in effortful, systematic reasoning, arriving at opinions unbiased by those of their discussion partner.⁶²

It is important to distinguish between the motivation to arrive at an accurate judgment and the concern to be a certain kind of person, the kind who is 'an independent thinker.' While the accuracy motivation demonstrably improves decision-making performance, the concern to be an 'independent' kind of person, with its characterological slant, can distort behaviour through the influence of self-presentation. In one series of studies a sub-group of experimental subjects, whose desired selves were distinguished by an emphasis on forming opinions independently of what others think, were indeed more likely not to conform under conditions that elicited conformity from other, less independent-minded subjects. But the self-professed 'independents' took it too far, going out of their way not to conform -- even to the point of publicly reporting attitudes that differed from their private beliefs, if such public reports would promote the impression that they were independent.⁶³

Accountability to independent, external review is the institutional counterpart of the individual accuracy motivation. Well-governed institutions should structure incentives for individuals accordingly.⁶⁴ For example, in the wake of 20th-century scandals involving medical research with human subjects, the protection of subjects' rights and welfare is no longer left entirely to the personal character of scientific

investigators, however well-intentioned; under international guidelines and under the law of many countries, it is subject to prospective and ongoing independent review, prompting investigators to think more sensitively and systematically about the ethical aspects of their research.⁶⁵ Similarly, the empirical literature on decision-making can inform the design of better mechanisms to promote accountability on the part of corporate leaders.⁶⁶

IV. Conclusion: Virtue Reconsidered

For people who care about living up to their own values, the activity known to psychologists as ‘self-regulation’ – monitoring one’s movement toward or away from the ‘desired self’ – is practically inescapable. The psychological evidence suggests that self-regulation influences conduct partly through its engagement with the ongoing social demands of self-presentation, often through cognitive processes that are not accessible to first-person introspection or control.

The influence of these social demands may edify personal character, but may also disrupt it – even when exerted through relationships with the *same people* on both counts, and even if both you and those other people (let us assume for the sake of argument) have the right values in principle. When your internal sense of ethical accountability is pervasively influenced by your relationships with close associates whose values you know, admire, and share – along the lines of the ideal Aristotelean social world – you run the risk of living in an incubator for your own cognitive complacency.⁶⁷ Even if the shared values are actually right, what threatens to dull your deliberative edge is your comfortable familiarity with the attitudes held in common. Familiarity breeds a pattern of

overconfident, low-effort, intuitive decision-making, which may sometimes fail to track important moral considerations. Being such as to make the right choices when you *do* think through your actions does not guarantee that you *will* in fact think through all of the actions you ought to think through.

Add to this an explicitly characterological outlook on self-regulation and self-presentation, manifested in ordinary circumstances of judgment and decision-making. Such an outlook can exacerbate a common cognitive pre-commitment to credit yourself, in broadly characterological terms, with doing your best to be ‘a good person,’ thus biasing retrospective self-assessment in your favour, and preparing the ground for future repetitions of every unwitting moral misapprehension.

But there is room for hope. The evidence suggests that cognitive complacency is remediable to a significant extent by the motivation to reach judgments and decisions that are well-founded and will stand up to critical scrutiny. And it may be possible to modulate one’s emotional investment in the wish to be ‘a good person,’ at least enough to reduce the risk of overlooking one’s own fallibility.

To revisit our pivotal question about moral theory and scientific psychology (Section I above), what import do these inductive generalizations have for virtue ethics? Do they merely lengthen the (already long) diagnostic list of ways in which people are liable to fall short of the ideal, or do they indicate some constraint on viable content for the normative psychology of virtue?

The problem made salient by the empirical generalizations discussed above is a common, well-intentioned inflection of the first-person perspective on ethical life: framing the ethical commitments that anchor self-regulation primarily as ideals of

personal character. To be sure, ordinary notions of personal character and ‘being a good person’ are certainly not the equivalent of philosophical virtue ethics. Nonetheless, what philosophical virtue ethics recommends to the reflective agent is recognizable as an intellectually disciplined cousin (and in some cases a cultural progenitor) of the ordinary notions whose routine involvement in self-regulation invites the troubling outcomes discussed in Section III.B above. Orthodox virtue ethics, unlike moral theories in the consequentialist or Kantian families, grounds the criterion of right action in some conception of what an idealized yet fully human agent would do. This builds into first-person ethical practice, as fundamental and indispensable, a characterological orientation that is entirely discretionary for consequentialist or Kantian ethical practice.

Moreover, given that a motivation for accuracy can remedy cognitive complacency, consequentialism and Kantianism each provide a sharply defined target for that motivation to home in on. Of course, this in itself does not vindicate either type of theory as correct. And obviously, no credible moral theory can interpret itself with respect to real situations and problems. The work of deliberation remains for the agent to do. The point is that virtue ethics makes extremely ambitious cognitive demands on the aspiring agent, in that it burdens the accuracy motivation with so many disparate considerations to track and organize under the abstract idea of ‘what the virtuous person would do.’ In order to ease the cognitive burden to the point of offering a reliable guide to first-person ethical practice, advocates of virtue ethics may be well-advised to make the way more navigable between that abstraction and the everyday arrival at choices of action.

With its ideal of personally embodied practical wisdom, virtue ethics offers a

criterion of right action intended to register the importance of multiple goods and evils in a unified rational perspective. Although, strictly speaking, the theory does not prescribe cognitive shortcuts, recommending instead that agents engage in deliberation to whatever degree is necessary to track all the goods and evils in play, the imagined exemplary figure of the virtuous person lends itself to oversimplified, self-enhancing cognition. All the more so when the exemplary figure you imagine is a composite drawn partly from (aspects of) real people, with whom you have close emotional or collegial bonds, and who regularly affirm your own understanding of yourself as a good person. The cognitive shortcut of doing what you expect your admired associates to approve of, using them as stand-ins for ‘the virtuous person’ or ‘the person of practical wisdom,’ becomes particularly tempting. But that shortcut, simply because it is a shortcut substituted for systematic deliberation, will not reliably lead to the right choice of action.

In order for virtue ethics to offer reliable guidance for first-person ethical practice, it may need to introduce an explicit shift of emphasis between the criterion of right action and the mode of deliberation recommended for everyday life. The criterion of right action is defined primarily in terms of what the virtuous person would do, and specified in detail this will encompass and unify a variety of substantive moral considerations. But how the aspiring agent ought to deliberate may be better articulated by some sort of explicit procedure for systematically identifying and thinking through those substantive considerations, as right-making and wrong-making features of particular prospective actions. The mental contents that ought to come to mind -- those that enable us to home in on the target of the motivation for moral accuracy -- are not adequately captured in the summary notion of ‘what the virtuous person would do,’ which well-intentioned agents

all too readily bring to mind and use without bothering to include the specification necessary to flesh out the summary. Rather, the appropriate mental contents should, as experienced, approximate as closely as possible *how the virtuous person would think*, and this requires that the agent be attuned primarily and systematically to the substantive goods and evils themselves.

The idea of ‘what the virtuous person would do’ does not in itself suffice to guide the agent in tracking and organizing substantive moral considerations as the virtuous person would. Indeed, neither does that idea in combination with a list (also abstract) of the considerations to which, in general, the virtuous person would be sensitive. What the aspiring agent needs, in addition, is more explicit guidance for charting particular courses of action by reference to those considerations. Whereas for Aristotle himself, such guidance was to be provided by mentors who would take the aspiring agent under their wing and by friends who would form an audience of like-minded peers, the empirical evidence surveyed above suggests that the agent will need to refer as well to some guiding procedure that remains independent of his or her interpersonal relationships. To formulate such guidance is an especially demanding challenge for virtue ethicists, so far as they hold (for good reason, in my opinion, although I will not argue for it here) that the plural nature of important ethical considerations makes decision-making algorithms impossible.⁶⁸

In conclusion, what the evidence about interpersonal processes reveals is a psychological need for systematic decision-making guidance, if aspiring agents are to adopt virtue ethics as a reliable guide to first-person ethical practice. This psychological need does not necessarily constrain the normative content of the criterion for right action,

but it does place a serious and challenging demand on virtue ethics to articulate some form of stand-alone decision procedure, comparable to the procedures that consequentialist and Kantian moral theories can more straightforwardly provide.

NOTES

¹ Credit for significant financial, institutional, and collegial support in the production of this article is due to the Edmond J. Safra Foundation Center for Ethics at Harvard University, where I held a Faculty Fellowship in 2005-06. For thoughtful reading and discussion of an earlier draft, I thank Dennis F. Thompson and the Faculty Fellows' seminar, and especially Renee M. Jones for extensive discussion, exchange of ideas, and fresh insights into the psychological literature. Precious Eboigbe provided excellent research assistance at Harvard. Grateful acknowledgment for discussion and comments at various stages of writing is also due to the following individuals and groups: Neera Badhwar, Lorraine Besser-Jones, Alan Fuchs, Maximilian de Gaynesford, Eric Hutton, Mark Leary, and Gopal Sreenivasan; the Moral Psychology Research Group, especially John Doris, Gilbert Harman, and Walter Sinnott-Armstrong; the Department of Philosophy at Wake Forest University; and the audience at the Rocky Mountain Virtue Ethics Summit, especially Rosalind Hursthouse, Karen Stohr, Christine Swanton, and Kit Wellman. Some financial support for work at earlier stages was provided by a Summer Research Grant at the College of William and Mary.

² 1105a30-1105b1. All quotations from the text of Aristotle's *Nicomachean Ethics*

(hereafter, *NE* for short) will be in David Ross's translation, as revised by J.L. Ackrill & J.O. Urmson. *Nicomachean Ethics*, (Oxford: Oxford University Press, 1925).

³ A central instance is Rosalind Hursthouse, *On Virtue Ethics* (Oxford: Oxford University Press, 1999), p.136 and pp. 113-119. On the centrality of practical reasoning – choosing the right actions for the right reasons – to the classical Western conception of virtuous character, see Julia Annas, 'Virtue Ethics', in D. Copp (ed.), *The Oxford Handbook of Ethical Theory*, (Oxford: Oxford University Press, 2006), pp. 513-536.

⁴ Jonathan Webber, 'Virtue, Character, and Situation', *Journal of Moral Philosophy* 3 (2006): pp 193-213; John M. Doris and Stephen P. Stich, 'As a Matter of Fact: Empirical Perspectives on Ethics', in F. Jackson and M. Smith (eds.), *The Oxford Handbook of Contemporary Philosophy* Oxford: Oxford University Press, 2005), pp. 114-152; John M. Doris, 'Precis' and 'Replies: Evidence and Sensibility.'" *Philosophy and Phenomenological Research* 73 (2005), pp. 632-635, 656-677, with commentaries by Julia Annas (pp. 636-642), Nomy Arpaly (pp. 643-647), and Robert C. Solomon (pp. 648-655); John Sabini and Maury Silver, 'Lack of Character? Situationism Critiqued', *Ethics* 115 (2005), pp. 535-562; Peter B. M. Vranas, 'The Indeterminacy Paradox: Character Evaluations and Human Psychology', *Nous* 39 (2005), pp. 1-42; Rachana Kamtekar, 'Situationism and Virtue Ethics on the Content of Our Character', *Ethics*, 114(2004), pp. 458-491; Peter Railton, 'Toward an Ethics that Inhabits the World', in Brian Leiter (ed.), *The Future for Philosophy* (Oxford: Oxford University Press, 2004), pp. 265-284; Shaun Nichols, *Sentimental Rules: On the Natural Foundations of Moral*

Judgment (Oxford: Oxford University Press, 2004); Christian Miller, 'Social Psychology and Virtue Ethics', *Journal of Ethics* 7 (2004), pp. 365-92; John M. Doris, *Lack of Character: Personality and Moral Behaviour* (Cambridge: Cambridge University Press, 2002); Gopal Sreenivasan, 'Errors about Errors: Virtue Theory and Trait Attribution', *Mind* 111 (2002), pp. 47-68; Maria Merritt, 'Virtue Ethics and Situationist Personality Psychology', *Ethical Theory and Moral Practice* 3 (2000), pp. 365-83; Gilbert Harman, 'Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error', in G. Harman, *Explaining Value and Other Essays in Moral Philosophy* (Oxford: Oxford University Press, 2000), pp. 165-78; Owen Flanagan, *Varieties of Moral Personality: Ethics and Psychological Realism*, (Cambridge: Harvard University Press, 1991).

⁵ Thanks to Geoff Sayre-McCord for conversation on this point. For a similar point based on common philosophical acceptance of the principle that moral judgments supervene on non-moral facts, see Railton, 'Toward an Ethics that Inhabits the World', pp. 270-271): "...if we act rightly or wrongly, for the good or the bad, we do so within the natural world we inhabit as empirical beings....any account of the domain of moral thought and practice must be compatible with what we know of the domain of human psychology, biology, and circumstance upon which it supervenes."

⁶ Kamtekar, 'Situationism and Virtue Ethics on the Content of Our Character', pp. 482-485.

⁷ Doris, *Lack of Character*, pp. 110-112.

⁸ Eric L. Hutton, 'Character, Situationism, and Early Confucian Thought', *Philosophical Studies* 127 (2006), pp. 37-58, at p. 49.

⁹ A kind of informal, non-institutional moral education may occur through young people's attempts to imitate morally exceptional community members whom they admire. But the uncritical acceptance of personal acquaintances as exemplars becomes problematic to the extent that moral maturity demands independent thought. (See Section IIIB below.)

¹⁰ Hutton, 'Character, Situationism, and Early Confucian Thought', p. 50. Hutton emphasizes that he presents this solution as an exegesis of the early Confucian thinkers, not as a positive practical proposal.

¹¹ Or third, as Hutton ('Character, Situationism, and Early Confucian Thought', p. 57, n. 41) suggests in passing, "insofar as a society is democratic, and it is the ordinary people who direct the government and thereby their own lives, it may seem equally imperative to make their behaviour as little situationally-dependent as possible." But to take this third option is to back off from the assumption that virtue is rare, thus incurring greater exposure of moral norms to empirical generalizations about what most people are like.

¹² In principle, my point about interpersonal processes should stand on its own, whatever

one's view of the situationist problem.

¹³ Cf. Bernard Williams, *Ethics and the Limits of Philosophy* (Cambridge: Harvard University Press, 1985), pp. 8-11.

¹⁴ *NE* 1106b36

¹⁵ While choosing the right actions may often be effortless for the virtuous agent, it is not mindless. As opposed to doing the same thing always and inattentively in the same circumstances, like brushing your teeth every night before bed, virtuous choice “involves doing the appropriately different thing attentively in varying circumstances.” W.F.R. Hardie, *Aristotle's Ethical Theory* (Oxford: Oxford University Press, 1968), p. 104.

¹⁶ *NE* 1114b27-29

¹⁷ *NE* 1140b29-30

¹⁸ *NE* 1104b30-1105a1. For Aristotle, the noble is associated with the idea of a given action as the done thing, or more precisely, what the man of practical wisdom would do.

¹⁹ *Ibid.*

²⁰ *NE* 1098b33-1099a5; see also *NE* 1114a7-10.

²¹ NE 1116a20-26

²² NE 1095b26-29

²³ See also NE 1159a21-24. For a discussion of the general point see Bernard Williams, *Shame and Necessity* (Berkeley: University of California Press, 1993).

²⁴ NE 1172a11-14

²⁵ NE 1104b30-1105a1

²⁶ Agnieszka Jaworska, 'Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value', *Philosophy and Public Affairs* 28 (1999), pp. 105-138, at p. 115.

²⁷ NE 1106b36-1107a2

²⁸ Hursthouse, *On Virtue Ethics*, p. 28.

²⁹ C.S. Carver, 'Self-Regulation', in A. Tesser and N. Schwarz (eds.), *Blackwell Handbook of Social Psychology: Intraindividual Processes* (Oxford: Blackwell, 2001), pp.307-328. See also C.S. Carver and M.F. Scheier, *On the Self-Regulation of Behaviour* (Cambridge: Cambridge University Press, 1998).

³⁰ E.T. Higgins, 'Self-discrepancy Theory: A Theory Relating Self and Affect', *Psychological Review*, 94 (1987), pp. 319-340; E.T. Higgins and O. Tykocinski, 'Self-discrepancies and Biographical Memory: Personality and Cognition at the Level of Psychological Situation', *Personality and Social Psychology Bulletin*, 18 (1992), pp. 527-535

³¹ For reviews of the literature, see: S. Chen, H.C. Boucher, and M.P. Tappin, 'The Relational Self Revealed: Integrative Conceptualization and Implications for Interpersonal Life', *Psychological Bulletin* 132 (2006), pp. 151-179; M.R. Banaji and D.A. Prentice, 'The Self in Social Contexts', *Annual Review of Psychology* 45 (1994), pp. 297-332; H. Markus and S. Cross, 'The Interpersonal Self,' in L.A. Pervin (ed.), *Handbook of Personality* (New York: Guilford Press, 1990), pp. 576-608.

³² M.R. Leary, 'The Self We Know and the Self We Show: Self-esteem, Self-presentation, and the Maintenance of Interpersonal Relationships', in G.J.O. Fletcher and M.S. Clark (eds.), *Blackwell Handbook of Social Psychology: Interpersonal Processes* (Oxford: Blackwell, 2001), pp. 457-477.

³³ In the routines of everyday activity and interaction, there is "simply too much information about oneself to do anything but provide an edited, packaged version that is relevant to the goals at hand." B.R. Schlenker and M.F. Weigold, 'Self-Consciousness and Self-Presentation: Being Autonomous Versus Appearing Autonomous', *Journal of Personality and Social Psychology*, 59 (1990), pp. 820-828. See also B.R. Schlenker and

M.F. Weigold, 'Interpersonal Processes Involving Impression Regulation and Management', *Annual Review of Psychology* 43 (1992), pp. 133-168; M.R. Leary, *Self-presentation: impression management and interpersonal behaviour* (Boulder: Westview Press, 1996).

³⁴ Leary, 'The Self We Know and the Self We Show', pp. 462-463.

³⁵ To be sure, even when there is some incentive to satisfy such expectations, there might also be other incentives to opt out. The advisability of opting out will depend in part on the cost of severing the relationships in play.

³⁶ S.M. Drigotas, S. W. Whitton, C.E. Rusbult, and J. Wieselquist, 'Close Partner as Sculptor of the Ideal Self: Behavioural Affirmation and the Michelangelo Phenomenon', *Journal of Personality and Social Psychology* 77 (1999), pp. 293-323.

³⁷ The more superficial form of reinforcement is emphasized by Kamtekar, 'Situationism and Virtue Ethics on the Content of Our Character', pp. 488-491.

³⁸ James B. Stewart, 'The Kona Files', *The New Yorker*, February 19 & 26, 2007, pp. 152-167.

³⁹ Stewart, 'The Kona Files', p 159.

⁴⁰ A judge later dropped all charges against Dunn. But by then the episode had imposed a heavy cost in organizational turmoil, as well as an enormous personal toll upon Dunn herself. Benjamin Pimentel, 'Court dismisses charges against former HP chair', *The San Francisco Chronicle*, March 15, 2007, p. C1.

⁴¹ H.M. Bazerman and M.R. Banaji, 'The Social Psychology of Ordinary Ethical Failures', *Social Justice Research* 17(2004), pp. 111-115; M.R. Banaji, 'Ordinary Prejudice', *Psychological Science Agenda* 14 (American Psychological Association, 2001), pp. 8-11.

⁴² D.A. Moore and G. Loewenstein, 'Self-Interest, Automaticity, and the Psychology of Conflict of Interest', *Social Justice Research*, 17(2004), pp. 189-202; N. Dasgupta, 'Implicit Ingroup Favouritism, Outgroup Favouritism, and their Behavioural Manifestations', *Social Justice Research*, 17(2004), pp.143-169; L.A. Rudman, 'Social Justice in Our Minds, Homes, and Society: the Nature, Causes, and Consequences of Implicit Bias', *Social Justice Research* 17(2004), pp. 129-142.

⁴³ Jason Dana and George Loewenstein, 'A Social Science Perspective on Gifts to Physicians from Industry', *JAMA* 290 (2003): pp. 252-255.

⁴⁴ Bazerman and Banaji, 'The Social Psychology of Ordinary Ethical Failures'.

⁴⁵ Daniel Kahneman, 'A Perspective on Judgment and Choice: Mapping Bounded

Rationality', *American Psychologist*, 58(2003), pp. 697-720

⁴⁶ Kahneman 'A Perspective on Judgment and Choice', p. 699.

⁴⁷ Alongside Kahneman's research program, another school of thought prominent in the empirical literature on intuition, judgment, and decision-making is led by Gerd Gigerenzer. See, e.g., Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group, *Simple Heuristics That Make Us Smart* (New York: Oxford University Press, 1999). Gigerenzer has identified a form of intuitive decision-making, "one-reason decision-making" which is ubiquitous in everyday reasoning and may often be useful, but which might occasionally lead to moral misapprehension. When there are too many relevant considerations to take into account in a limited amount of time, we recruit a single type of consideration to stand in for the harder-to-determine, all-things-considered best choice. See also Walter Sinnott-Armstrong, 'Moral Intuitions as Heuristics', unpublished manuscript.

⁴⁸ Returning to the real-life example mentioned above, Patricia Dunn has reportedly said, "There has never been a whiff of scandal or taint related to my activities, particularly any issues concerning my integrity or ethics. Indeed, in the roles I have held, any such taint would be an instant career-ender, and for good reason." Quoted in Stewart, 'The Kona Files', p. 154. My point in quoting Dunn here is not to accuse her of moral conceit, but rather to note that even seemingly well-founded confidence in the quality of one's character can sometimes be ethically counterproductive.

⁴⁹ John Darley, 'The Cognitive and Social Psychology of Contagious Organizational Corruption,' *Brooklyn Law Review*, 70 (2005), pp. 1177-1194; Elliot Aronson, *The Social Animal*, 9th edn (New York: Worth Publishers, 2003), pp. 143-199; R.M. Jones, 'Law, Norms, and the Breakdown of the Board: Promoting Accountability in Corporate Governance', *Iowa Law Review*, 92 (2006), pp. 105-158.

⁵⁰ G. Weary, 'Self-presentation and the moderation of self-serving attributional biases', *Social Cognition* 1 (1982), pp. 140-159; Aronson, *The Social Animal*, p. 140; Jones, 'Law, Norms and the Breakdown of the Board'.

⁵¹ A.G. Greenwald and S.J. Breckler, 'To Whom Is the Self Presented?', in B.R. Schlenker (ed.), *The Self and Social Life* (New York: McGraw-Hill, 1985), pp.126-145, quoted in Aronson, *The Social Animal*, p. 139.

⁵² Kahneman, 'A Perspective on Judgment and Choice', p. 699.

⁵³ Kahneman, 'A Perspective on Judgment and Choice', p. 698. See also Moore and Loewenstein, 'Self-Interest, Automaticity, and the Psychology of Conflict of Interest'.

⁵⁴ Kahneman, 'A Perspective on Judgment and Choice', p. 698. It is important, however, not to over-generalize about the type of cognitive processing typically associated with moral judgment. Even as many aspects of moral judgment draw on slow, deliberative cognitive processing, evidence suggests that other aspects are sub-served by fast,

automatic cognitive processing. See, *e.g.*, Joshua D. Greene, *et al.*, 'The neural bases of cognitive conflict and control in moral judgment', *Neuron* 44 (2004), pp. 389-400.

⁵⁵ Darley, 'The Cognitive and Social Psychology of Contagious Organizational Corruption', p. 1185.

⁵⁶ J.S. Lerner and P.E. Tetlock, 'Accounting for the Effects of Accountability,' *Psychological Bulletin*, 125 (1999), pp. 255-275, at p. 255.

⁵⁷ Lerner and Tetlock, in 'Accounting for the Effects of Accountability', provide a comprehensive review and analysis of the extensive body of research on accountability and its impact on judgment and decision-making.

⁵⁸ Lerner and Tetlock, 'Accounting for the Effects of Accountability'; J. Pennington, J. and B.R. Schlenker, 'Accountability for Consequential Decisions: Justifying Ethical Judgments to Audiences', *Personality and Social Psychology Bulletin*, 25(1990), pp. 1067-1081.

⁵⁹ Lerner and Tetlock, 'Accounting for the Effects of Accountability'; see especially Table 1, p. 260.

⁶⁰ S. Chen, D. Schechter, and S. Chaiken, 'Getting at the Truth or Getting Along: Accuracy-Versus Impression-Motivated Heuristic and Systematic Processing', *Journal of Personality and Social Psychology*, 71(1996), pp. 262-275; A. Quinn, and B.R.

Schlenker, 'Can Accountability Produce Independence? Goals as Determinants of the Impact of Accountability on Conformity', *Personality and Social Psychology Bulletin*, 28 (2002), pp. 472-483. See also Jones, 'Law, Norms and the Breakdown of the Board'.

⁶¹ Quinn and Schlenker, 'Can Accountability Produce Independence?'

⁶² Chen, Schechter, & Chaiken, 'Getting at the Truth or Getting Along'.

⁶³ B.R. Schlenker and M.F. Weigold, 'Self-Consciousness and Self-Presentation: Being Autonomous Versus Appearing Autonomous,' *Journal of Personality and Social Psychology* 59 (1990), pp. 820-828.

⁶⁴ However, for some domains, the empirical literature arguably suggests that self-interested biases operate so far below the radar of conscious, deliberate cognition that it may be more effective simply to remove the distorting influences altogether. For instance, if the evidence shows that neither the disclosure of physicians' involvement with pharmaceutical companies nor limits on the size of promotional gifts to physicians will work to correct bias in prescribing practices, a policy of outright prohibition might be preferable. Dana and Loewenstein, 'A Social Science Perspective on Gifts to Physicians from Industry', p. 254.

⁶⁵ Henry K. Beecher, 'Ethics and Clinical Research', *New England Journal of Medicine* 274 (1966), pp. 1354-1360; Allan M. Brandt, 'Racism and Research: The Case of the

Tuskegee Syphilis Study’, *Hastings Center Report* 8 (1978), pp. 21-29; National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, ‘The Belmont Report’, U.S. Government Printing Office, April 18, 1979; U.S. Code of Federal Regulations Title 45, part 46, Protection of Human Subjects; Council for International Organizations of Medical Sciences (CIOMS), *International Ethical Guidelines for Biomedical Research Involving Human Subjects* (Geneva: World Health Organization 2002).

⁶⁶ Jones, ‘Law, Norms and the Breakdown of the Board’.

⁶⁷ Contrast the contemporary ideal of social diversity in educational environments.

⁶⁸ Cf. Hursthouse, *On Virtue Ethics*; Christine Swanton, *Virtue Ethics: A Pluralistic View*. (Oxford: Oxford University Press, 2003).