

## CONTENTS

*Preface* ix

|                      |   |     |
|----------------------|---|-----|
| 1                    | Credo   | 1   |
| 2                    | The Integral View                                 | 22  |
| <b>PRELIMINARIES</b> |   |     |
| 3                    | Emotion   | 37  |
| 4                    | Disunity of Virtue                                | 67  |
| 5                    | Character Traits                                  | 84  |
| <b>ARGUMENTS</b>     |   |     |
| 6                    | Adverbial Requirements                            | 111 |
| 7                    | Salience without a Black Box                      | 129 |
| 8                    | Moral Deference and the Proto-authority of Affect | 168 |
| 9                    | Recap with Courage                                | 199 |
| <b>CONSEQUENCES</b>  |   |     |
| 10                   | Agents versus Acts                                | 243 |
| 11                   | Against the Priority of Principle                 | 266 |
| 12                   | <i>Should</i> Virtue Be Taught?                   | 299 |

viii CONTENTS

*Acknowledgements* 327

*Notes* 331

*References* 375

*Index* 387

# 1

## Credo

Credo virtutem nec unam nec omnipotentam

§1. WHAT IS A VIRTUE? There are several different ways to hear this question. On perhaps the most basic way of hearing it, a long-standing philosophical tradition answers that a moral virtue is a species of character trait. While some contemporary philosophers depart from this tradition, even they do not deny that virtuous character traits and virtuous actions can be inter-defined. Indeed, they can be inter-defined in various, verbally equivalent ways. For example, a kind person can be defined as someone who, among other things, can be relied upon to act kindly—to help an old woman to cross the street, say, or to overlook faults in others.\* Alternatively, kind actions can equally be defined as the sort of actions characteristically performed by a kind person (i.e., a person with a certain trait). In one sense, then, kindness can easily be regarded as both a character trait and a species of action.

To isolate the feature of the traditional view in dispute, it therefore helps to advert explicitly to the *direction of priority* being affirmed when virtuous traits and virtuous actions are inter-defined. Are kind actions basic, with the character trait of kindness defined derivatively (in terms of them)? Or is the priority rather reversed, with the character trait being basic and kind actions being defined derivatively (as characteristic expressions of the trait)? On the traditional view, character traits have priority in the definition of virtue.

By contrast, Thomas Hurka (2006) defends an ‘occurrent-state view,’ according to which virtuous actions have priority. Specifically, on his view, an action is virtuous if and only if it is virtuously motivated; and independent conditions are given to define what makes a motive or desire virtuous. Crucially, these conditions do not refer to any disposition or character trait.

\* Throughout this chapter, kindness stands in for any old specific virtue.

Thus, an agent's occurrent desire (and hence her occurrent action when so motivated) can satisfy these conditions, and thereby qualify as virtuous, even if it is 'out of character' or is a onetime occurrence. Virtuous character traits are then defined derivatively, as dispositions to perform virtuous actions.

Hurka opposes his view to a 'dispositional' view. However, as he defines it, the dispositional view does more than merely reverse the direction of priority from the occurrent-state view. For it not only defines a virtuous act derivatively, as a characteristic expression of a virtuous disposition, but *also* requires the agent performing the act *herself to possess* the virtuous disposition in question: 'the dispositional view also holds that virtuous [occurrent] states necessarily issue from virtuous dispositions' (Hurka 2006: 71). Consequently, if an agent lacks the trait of kindness, it follows that no act that *she* performs can count as a kind act—however much it may otherwise resemble the acts characteristically performed by kind people.

Hurka rejects this implication, and rightly so. Yet it is important to see that traditional views on the definition of virtue need not endorse the implication he rejects, since they need not embrace the additional requirement built into his dispositional view. Consider, for example, Aristotle's well-known distinction between performing a virtuous act and performing a virtuous act *as* an exemplar of virtue would perform it (*EN* 1105b5–9), i.e., as a model or paragon of virtue would perform it. Exemplars of virtue have, and characteristically act from, a stable disposition. *A fortiori*, no one can act 'as an exemplar of virtue would' without herself acting from a stable disposition.<sup>1</sup> However, the whole point of Aristotle's distinction is precisely to allow that someone can still perform a virtuous act—can still do *what* the exemplar of virtue does—even if he or she cannot perform that act *as* the exemplar performs it (e.g., because he or she lacks the relevant stable disposition).

Accepting Aristotle's allowance is entirely compatible with reversing the direction of priority from the occurrent-state view. It is compatible, that is, with insisting that what nevertheless *makes* the act in question 'virtuous' is that it is the characteristic expression of a certain trait, i.e., that it is what an exemplar of virtue would characteristically do (under the circumstances). For example, what makes 'helping an old woman to cross the street' a *kind* act is that helping her to cross is what an exemplar of kindness would characteristically do. On the resultant (perfectly traditional) view, the kindness of the act is derivatively defined, and this derivation refers to a character trait, but that prior trait need not be possessed *by* the agent who performs the act (in order for it to be a kind act). Let us distinguish this logically weaker opponent of the occurrent-state view from the dispositional view by calling it the (metaphysi-

cal) *agent-centred* view.\* Since Aristotle's distinction is what opens the door to a perspicuous statement of this weaker view, it seems a mistake to saddle *him* with Hurka's dispositional view (and its implausible implication).<sup>2</sup>

Like Hurka (2006), Judith Thomson (1997) also rejects the traditional direction of priority in the definition of virtue. However, unlike Hurka, Thomson construes occurrent virtuous acts strictly *objectively* (1997: 281, 286). Hence, in addition to excluding reference to any disposition or character trait, the conditions her view employs to define what makes an act virtuous also exclude reference to the agent's occurrent motive or intention.

One way to understand this narrowed scope of the 'occurrent acts' to which Thomson gives metaphysical priority is to see her as shifting the agent's occurrent motive from the 'performs a virtuous action' side of Aristotle's distinction to its 'performs a virtuous action *as* the exemplar of virtue performs it' side (cf. endnote 1), where what remains on the first side still suffices to qualify an occurrent performance as a virtuous act.<sup>3</sup> Thus, someone who helps an old woman to cross the street, even from an unsuitable motive, still performs a kind act. That is, he still does *what* an exemplar of kindness does. Naturally, he deserves less (and perhaps very little) credit for doing so, but that is another matter.<sup>4</sup>

I believe that, metaphysically, kind actions are basic and that the virtuous character trait of kindness should be defined derivatively, as a disposition to perform kind actions (among other things). Let us call this the metaphysical act-centred view of virtue. Rather than argue for this position, I shall take it as intuitive.

To this point, I have discussed the priority question in metaphysical terms, since that is the predominant treatment in the literature. But let us now proceed to distinguish metaphysical and epistemological versions of the priority question. A sufficient reason to do so is that one may wish to answer the two versions of the question differently (e.g., I do). In the epistemological case, the priority question concerns the starting point for identifications of virtue. Do we first identify a character trait as virtuous (or, more specifically, as kind), and only then identify its characteristic act expressions as virtuous acts (or kind acts)? Or do we rather identify various acts first as kind acts, and only then identify the persons who reliably perform those acts as kind persons (i.e., as having that trait)?

\* Among contemporary philosophers, Gary Watson (1990) and Michael Slote (1997) defend the metaphysical agent-centred view.

According to the *epistemological act-centred* view of virtue, the priority for identifying instances of virtue lies with virtuous acts, whereas according to the *epistemological agent-centred* view it lies with virtuous traits instead.<sup>5</sup> Unlike with the metaphysical priority question, I do not believe that the act-centred view gives the correct answer here. To make a start on seeing why not, we should notice that the rival answers to both priority questions have been presented in all-or-nothing terms. In the metaphysical case, this arguably makes ready sense. But in the epistemological case, it obscures a coherent intermediate position.

Consider the following entailments of the two extreme answers to the epistemological priority question: On the act-centred view, it follows that *every* kind act, say, can be identified as kind without any reliance on a kind person.\* On the agent-centred view, by contrast, it follows that *no* kind act can be identified as kind without some reliance on a kind person. Evidently, there remains the intermediate possibility that *some* kind acts can be identified as kind without any reliance on a kind person, while other kind acts *cannot* be so identified except by relying somehow on a kind person.

Whatever else one thinks of it, this intermediate option is immune to the two most obvious objections against the agent-centred view.† One obvious objection is that, intuitively, some kind acts *are* straightforwardly identifiable as kind without one's either being or referring to a kind person. 'Paradigmatic' or 'stereotypical' acts of kindness, such as helping an old woman across the street, seem to have this feature by definition. Another popular objection claims that the agent-centred view cannot explain how anyone can non-arbitrarily identify who the kind persons are (e.g., O'Neill 1996; Cholbi 2007). Or, more generally, it cannot identify who the virtuous persons are.

Let me therefore introduce the *modest agent-centred* view. According to this intermediate answer to the epistemological priority question, some non-paradigmatic acts of kindness (say) can only be identified as kind acts by exploiting the fact that they are the characteristic act expressions of a certain trait (kindness). More specifically, some kind acts cannot be identified as kind except by relying on a kind person one way or another. Since it is common ground between the modest agent-centred view and the act-centred view, the fact that paradigmatic acts of kindness can be identified as kind without rely-

\* One relies on a kind person to identify an act as kind if one either has to *be* a kind person oneself or has to *refer to* a kind person, in order to identify that act as kind.

† Henceforth, I shall omit the qualification 'epistemological,' which should be understood. In this book, my primary concern with the priority question is with its epistemological version.

ing on any kind person grounds no objection to the modest agent-centred view. Moreover, the same fact leaves it open to the modest agent-centred view to hold that kind persons are to be identified by means of their reliability in performing paradigmatically kind acts (a criterion that anyone can employ), thereby defusing the second objection.

I believe that the modest agent-centred view gives the correct answer to the epistemological priority question.\* Now a case can be made that this same view could equally well be called the 'modest act-centred' view. However, I shall nevertheless call it the modest *agent*-centred view. I do so in order to advertise the fact that it preserves an indispensable role for virtuous traits in the identification of virtuous actions, and thereby partially vindicates the traditional view of virtue. As I see it, the modest agent-centred view articulates what is correct in the traditional answer to the epistemological priority question. Where the (extreme) act-centred view goes wrong, by contrast, is in denying that virtuous traits have any indispensable role at all in identifying virtuous actions.

§2. If a virtue is a species of character trait, it is very natural to wonder further: what sort of character trait is it?<sup>†</sup> Or, to adopt somewhat different phrasing, what are the 'other things' referred to in the proposition that a kind person can be relied on to act kindly, among other things? Theories of virtue disagree about what else is required, for reasons having nothing to do with the particulars of kindness. By way of illustration, consider the following three possible additional requirements for virtue (we shall encounter others in time). All three are affirmed by Aristotle, for example.

One might require, first, that a kind person not only reliably does the kind thing, but also (reliably) does the kind thing *for the right reason(s)*. Second, one might require that a kind person (reliably) does the kind thing in some particular way, rather than in just any old way. For example, one might require that her reliably kind actions are performed *wholeheartedly*. Third, one might require that a kind person also (reliably) does the *just* thing, and the *brave* thing, and the *generous* thing, and so on. That is, one might affirm the 'unity,' or better, the 'reciprocity of the virtues,'<sup>‡</sup> understood as a requirement on a person's qualifying even simply as *kind*.

\* I provide a positive argument for the modest agent-centred view in chapter 10. For the most part, the intervening discussion will concentrate on paradigmatic acts of virtue. Under that constraint, the disagreement between the modest agent-centred view and the (extreme) act-centred view is less germane.

<sup>†</sup> This is another way to hear our opening question, what is a virtue?

<sup>‡</sup> This thesis holds that in order to have a given virtue, one must (also) have all of the other virtues. On the distinction between reciprocity and unity, see Irwin (1997).

I believe that the first two additional requirements belong to the nature of virtue (at least, some version of each does). By contrast, I reject the third requirement—hence the first half of my epigraph, *credo virtutem nec unam*. However, since I shall argue against the reciprocity of the virtues at some length in chapter 4, let me defer consideration of it until then. What I primarily wish to discuss here is the right reasons requirement. But let me begin by addressing the second additional requirement briefly.

In Book VII of his *Nicomachean Ethics*, Aristotle sets out a tripartite typology of ethical characters, distinguishing the virtuous person from both the ‘strong-willed’ person and the ‘weak-willed’ person. All three types reliably make the ethically correct choice, reliably choose the virtuous thing to do. The weak-willed person is distinguished from the other two by the fact that he cannot be relied upon to act on his virtuous choice consistently. As this suggests, the virtuous person and the strong-willed person actually have it in common that they consistently act on their virtuous choices. What distinguishes the strong-willed person from the virtuous person is that the strong-willed person is somehow conflicted about his virtuous choices, and must therefore overcome some internal obstacle(s) in order to act on them. Nevertheless, precisely because he is *reliably* strong-willed he does consistently overcome whichever of his desires oppose the virtuous choice, and so can still be relied upon to do the virtuous thing consistently.

On this conception of virtue, it is a mark of virtue not to be internally conflicted about the virtuous thing to do. This is the mark I meant to capture with the requirement that a kind person’s reliably kind actions be performed ‘wholeheartedly.’ Of course, it is plainly easier—in one way, anyhow—to be reliable at doing the kind or virtuous thing if one does not face any internal obstacles to acting on one’s virtuous choice. In elevating virtue (so defined) above strength of will, Aristotle claims, in effect, that it is morally better for a person not to be beset by such internal obstacles in the first place, even if that makes the reliable performance of virtuous action easier. While his claim is certainly controversial, at least in contemporary discussion,\* I confess to a certain sympathy for it.

We shall examine the merit of Aristotle’s claim, up to a point, in chapter 9, in connection with a discussion of courage (a virtue for which his claim is probably least plausible). Our examination can be safely both postponed and limited because, despite my own sympathy for the distinction between virtue and strength of will, I shall nowhere rely on it. I invoke it here only to illustrate

\* Bernard Williams, for example, dismisses the distinction between virtue and strength of will out of hand as ‘a tedious Aristotelian ideal’ (1995: 194).

what I shall call ‘adverbial requirements’ on virtue—requirements that a person must perform the virtuous action *in a certain way* to qualify herself as virtuous—and I chose this illustration simply because it is the most famous. Other adverbial requirements will play a more prominent role in my argument. But, as we shall see, they will be far less controversial.

Let us now return to the right reasons requirement. Like adverbial requirements on virtue, the requirement that a kind person, say, reliably do the kind thing ‘for the right reason(s)’ is better seen as a family of requirements.<sup>6</sup> In particular, we should understand it as comprehending at least three distinct, though related, kinds of requirement on virtue. Two of them, I take it, are uncontroversial, while within bounds we can remain open-minded about the third.

To begin with, ‘right reasons’ can be read as reminding us that, even as a necessary condition on virtue, the basic idea that a kind person reliably does the kind thing is in one important respect oversimple. It would be more precise to say, with John McDowell,<sup>7</sup> that a kind person reliably does the kind thing *when that is what the situation requires*.<sup>\*</sup> A kind person does not help an old lady to cross the street when the street is full of traffic nor when the old lady does not want to cross. Even paradigmatic behavioural expressions of kindness, that is, cannot be understood as mechanically required by kindness. Rather, ‘kind’ behaviour, paradigmatic or not, is only required by kindness when, minimally, it will achieve some good—‘helping’ an old lady to cross the street, for example, is only required when crossing will actually help *her*. Hence, the judgement that a particular act *is* an act of kindness (or, more generally, of virtue) always depends upon some contextual evaluation. While some might regard this qualification as already implicit in the expression ‘act of kindness,’ the important further point is that the relevant evaluative dependence is (equally) a characteristic feature of a kind person’s reliability. As McDowell puts it, ‘a kind person has a reliable sensitivity to a certain sort of requirement which situations impose on behaviour’ (1979: 51).

A second component of the right reasons requirement is purely negative: to require that a kind person (reliably) does the kind thing for the right reasons is, minimally, to exclude her doing the kind thing for the *wrong* reason. Wrong reasons for doing the kind or virtuous thing centrally include doing it to

<sup>\*</sup> This expression is actually ambiguous as between ‘when that is what *kindness* requires in the situation’ and ‘when that is what *morality* requires in the situation.’ The ambiguity becomes significant in cases where the all-things-considered verdict of morality overrules the requirements of kindness. These cases raise important difficulties for the theory of virtue, which we shall have occasion to examine in chapter 4 (in connection with the reciprocity of the virtues) and again in chapter 12. In this chapter, however, I ignore the ambiguity for the most part.

impress someone or to develop a certain reputation—more generally, any reason that might be intuitively regarded as ‘ulterior.’ I take it that this much is straightforward.

The third component of the right reasons requirement adds some positive characterisation of the agent’s reason(s) for performing the kind act. At a minimum, this will tie the agent’s reasons for acting, whatever they are, to a description under which the act *is* an act of kindness (in the situation).<sup>\*</sup> That is, the kind agent’s act must be intentional under some such description, since one cannot act virtuously by accident. At the other extreme from this minimal intentionality requirement lies the requirement that a kind agent’s reasons for performing the kind act explicitly include the reason ‘that it would be the kind thing to do’ or ‘that kindness requires it’ or some such. I shall simply take it for granted that a theory of virtue (can and) should reject this extreme version of a positively characterised right reasons requirement.<sup>8</sup>

Still, there are presumably also intermediate possibilities to consider. Is there any positive characterisation more robust than a minimal intentionality requirement, yet less extreme than ‘kindness requires it’ that a kind agent’s reasons for performing the kind act has to satisfy? I shall describe two candidates and reject them both as otiose. But I am happy to leave the merits of other intermediate options as an open question.

One very natural possibility is to require that a virtuous agent’s intention in acting virtuously be a *good* intention. This option is somewhat controversial. Pure externalists about virtue, for instance, deny that there is any such requirement, even when ‘good intention’ merely means ‘not a bad intention.’<sup>9</sup> A fortiori, they deny that the virtuous agent must know (or even, believe) that her act is the ‘right’ act, in the sense of being consistent with the all-things-considered verdict of morality. Indeed, their main argument actually runs the other way round. It begins, that is, by appealing to cases in which agents (e.g., Huckleberry Finn) are said to act virtuously, despite themselves believing that the acts they perform are wrong; and then concludes on that basis that good intentions are not necessary for virtue.

The externalist argument requires us to accept that an agent’s being reliable at doing the kind thing (say) is compatible with her *believing* that (some of) her kind actions are inconsistent with the all-things-considered verdict of morality (henceforth, ‘ATC wrong’). In the scenario on which externalists focus, the agent has radically mistaken beliefs about the true requirements of morality (e.g., she believes in slavery): nevertheless, by stipulation, her kind actions

<sup>\*</sup> Thus, e.g., the agent must intend ‘to get the old lady across the street,’ rather than ‘to create a hazard in the road’ or ‘to add to the population on the opposite sidewalk.’

are actually perfectly consistent with those requirements. Thus, we are importantly not asked to count any actions that are actually ATC wrong as ‘kind’ or ‘virtuous’ actions (nor to count agents who perform such actions as ‘virtuous’ agents). However, this is not the only relevant scenario.

In particular, we should consider the scenario in which the agent is *correct* to believe that some of her kind actions are ATC wrong.\* Here we cannot avoid the question of whether ATC wrong actions are properly counted as ‘kind’ or ‘virtuous.’ The conventional answer to this question is a resounding ‘no’; and externalists themselves accede to convention on this point.<sup>10</sup> Let me accordingly follow suit for present purposes.<sup>†</sup> To accommodate this assumption, we need to introduce a distinction and a clarification. Thereafter, as we shall see, things unravel quickly for the externalist’s position.

Strictly speaking, the expression ‘kind actions that are ATC wrong’ now fails to refer. We therefore need to distinguish kind actions from ‘*superficially* kind’ actions, where the latter may be understood as actions that are required by kindness, *other things being equal*. Unlike (genuinely) kind actions, superficially kind actions can be ATC wrong. Yet when they are—and this is the clarification—they then function like actions that are ‘unkind,’ rather than merely ‘not kind.’ That is, performing them counts *against* an agent’s reliability as a kind agent.

It follows that a reliably kind agent cannot perform superficially kind actions that are ATC wrong.<sup>‡</sup> Alternatively, if she does, she no longer counts as reliably kind. But, of course, this same point equally limits what such an agent can correctly believe about *her own* actions, since an action’s being ‘hers’ implies that she performed it. Thus, no reliably kind agent can correctly believe that her own superficially kind actions are ATC wrong. When the agent’s beliefs are correct, then, the externalist’s claim that an agent’s being reliably kind

\* To keep things interesting, we have to assume that the scope of ‘some’ here *exceeds* the permitted margin of unreliability, whatever that might be. Otherwise the whole phenomenon of interest will just be lost in the margin of error. So if virtuous agents are required, e.g., to be 90 percent reliable, the correct belief in the text should be read as ‘more than 10 percent of her kind actions are ATC wrong.’

† Since I reject the reciprocity of the virtues (in chapter 4), I should say explicitly that the conventional position here is closely related to that thesis. Consequently, the argument to follow in the text is rather weaker under premisses that I myself accept. If one rejects the reciprocity of the virtues root and branch, as I do not, the argument does not work at all. Yet, as we shall also see, neither the argument nor its conclusion is crucial for my purposes.

‡ Throughout this paragraph, I assume that we are again discussing performance rates that exceed the margin of unreliability.

is compatible with her believing that some of her (superficially) kind actions are ATC wrong is simply false.<sup>11</sup>

This conclusion brings us rather closer to the proposition that a kind agent's intention in acting kindly must be a good intention. For not only must her act be intentional under some description that makes it an act of kindness in the situation, but it is now a condition of that's being possible—that is, a condition of any description's making her act one of kindness—that the act be morally permissible, all things considered. Hence, the kind agent's intention in acting kindly must respond to the (balance of considerations that determine the) overall permissibility of her act.

I do not know whether it is possible for actual human beings to be *reliable* at acting in line with the overall verdict of morality *without* aiming, directly or indirectly,<sup>12</sup> to obey that verdict.\* If it is, then there is room, even outside the margin of unreliability, for a kind agent's intention in acting kindly to be a neutral intention, rather than a good one. But if it is not possible—and I incline to doubt that it is possible—then having a good intention will turn out to be part and parcel of a kind agent's reliability in acting kindly. In that case, there is no need to ground the goodness of a kind agent's intention in acting kindly in any additional, free-standing requirement on virtue, since it is secured by the reliability requirement that is already in place.

A second possibility, related to the first but less demanding, is to require that a kind agent's *motive* in acting kindly be a good motive.<sup>†</sup> However, considerations very similar to those marshalled by the previous argument show that having a good motive is also part and parcel of a reliably kind agent's reliability in acting kindly.<sup>13</sup> In both cases, the basic question is whether the actions of actual flesh and blood human beings—as opposed to the stipulative creatures of thought experiments—can reliably achieve a certain moral status (here, [superficial] kindness; there, ATC permissibility) if the agent does not somehow *aim* to achieve that status. Each argument claims, as its central premiss, that the answer to this question is 'no.'

With respect to this central premiss, the argument is actually in two ways stronger when deployed against the good motive requirement, as compared to its good intention counterpart. To begin with, precisely because an agent's motive can be good even if her act is ATC wrong, the present argument can

\* Externalists evade this question by stipulating the outcome in the text. Of course, the answer will also depend on *how* reliable we understand 'reliable' to be (see further §3).

† This option is less demanding than the previous one insofar as the goodness of an agent's good motive in acting (superficially) kindly on some occasion is independent of whether her (superficially) kind act is ATC wrong. Thus, her motive can be good even if her act is ATC wrong.

allow that *kind* acts—or, more generally, virtuous acts—can be ATC wrong, thereby eschewing any commitment to the reciprocity of the virtues. We can therefore also dispense with the distinction between kindness and superficial kindness.

In addition, the case for the central premiss itself is even more compelling here: On any given occasion—or even, scattering of occasions—it is perfectly plausible that an agent might perform a kindness intentionally, but do so from a morally neutral motive. (She might intend, that is, to do something under a description—getting an old lady across the street, say—that makes her act one of kindness in the situation.) However, to perform kind actions *reliably* is another matter. Given that the identity conditions for a kind action depend on contextual evaluation, it seems there are really only two routes by which human beings can be reliably kind. Either a person explicitly aims to act kindly or under some equivalent description (and reliably succeeds) *or* she reliably responds more immediately to the underlying kindness facts across various situations,\* but without acting consistently under any particular description.

On both routes the agent arguably ‘aims’ in some sense to act kindly.<sup>†</sup> But we need not insist on that description with the second route. For it is actually more obvious that the mechanism invoked there—a person’s motives being reliably (positively) responsive to the underlying kindness facts—itself suffices to make the relevant motives *good*. It does not matter whether the person’s motives respond to the underlying kindness facts immediately or only mediately (by reliably subserving something else suitably responsive).<sup>14</sup> Either way, this route to reliability wears the desired conclusion more or less on its sleeve: having a good motive is part and parcel of reliability in acting kindly.

A further step, however, is required to reach that conclusion on the first route. For explicitly aiming to act kindly is, of course, fully consistent with having bad motives; and reliably succeeding in that aim does not alter this fact. Fortunately, the missing step is simple to supply. We only have to recall that the second component of the right reasons requirement has already excluded doing the kind thing for the wrong reason. In that case, what remains is actually a variant of the previous mediated reliability scenario, where the person’s at-worst-neutral motives reliably subserve something—here, her explicit

\* That is, she reliably responds, across various situations, to the considerations that determine the requirements of kindness. Of course, reliability may also be achieved by a mixture of the two routes described in the text.

<sup>†</sup> On the second route, the person’s aiming can, if you like, be understood as aiming *de re*, rather than *de dicto*.

aims—that reliably responds to the underlying kindness facts. Yet, as we said, this suffices to makes her motives (minimally) good.

Like her good intention, then, the goodness of a kind agent's motive in acting kindly is also part and parcel of her reliability in acting kindly (cf. end-note 13). So, once again, there is no need to ground the goodness of her reasons for acting kindly in any additional, free-standing requirement on virtue.

§3. Reliability in doing the kind thing is necessary for the virtue of kindness, though not sufficient (not even when the kind thing is reliably done for the right reason). But reliability itself is, of course, a matter of degree. This means that the possession of kindness—and, more generally, of virtue—lies on a continuum. While the continuum ranges, in principle, from zero to one hundred percent reliability, two points on it are of greater interest than the others. Both points function as thresholds. At the lower end of the continuum, there is a threshold for qualifying as (even) 'minimally kind' and, at the higher end, there is another threshold for qualifying as an 'exemplar of kindness.'\*

The minimal threshold for kindness marks the point below which a person does not even count as 'approximately kind' or as 'somewhat kind.' Rather, he or she is at best someone who occasionally does the kind thing. I doubt there is any way to remove all traces of arbitrariness from this distinction. However, the element of arbitrariness would be heavily reduced here if the goodness of a reliably kind agent's intention or motive in acting kindly were secured, as argued earlier, as a side effect of her reliability. For even if it obtains, this side effect presumably only kicks in above a certain point on the reliability continuum: very low levels of reliability in acting kindly, I take it, are fully consistent with an agent's having merely neutral intentions or motives. By the same token, once the idea of *degrees* of reliability has been made explicit, the proposition that the goodness of a kind agent's motive, say, in acting kindly is part and parcel of her 'reliability' in acting kindly becomes ever more plausible, the higher the degree of reliability one has in mind.

But then, returning to Aristotle's distinction between performing a virtuous act and performing a virtuous act *as* the virtuous person performs it, we can helpfully redescribe the point at which this side effect kicks in. Above the point at which good motives emerge on the reliability continuum, the reliably kind agent's kind actions are also (and therefore) performed *as* the kind per-

\* Tidiness would be most served if the two points of greatest interest actually coincided with the end points of the continuum—if zero were the 'threshold' for minimal kindness and exceptionless reliability were the 'threshold' for exemplar status. Nothing that I really want to say depends on rejecting this tidy view, but it strikes me as highly implausible.

son performs them—at least in the central respect of being performed with a good motive. Moreover, this holds true even if the agent’s reliability in acting kindly otherwise leaves much to be desired. Below this point on the continuum, by contrast, no one is licensed to infer that the low level reliably kind agent’s kind actions are performed with a good motive. Hence, while this person clearly does perform the occasional kind act, there is not the same basis for counting her as ‘kind’ in the thicker sense of performing kind acts *as* the kind person performs them. Accordingly, it makes ready sense to treat the point on the reliability continuum at which good motives emerge as the threshold for minimal kindness, understood in this thicker sense.

I believe that a minimal threshold for virtue is defensible. Nevertheless, the fundamental threshold on the reliability continuum is the higher one, the threshold that qualifies an agent—so far as reliability in the performance of the relevant acts is concerned—as an *exemplar* of a given virtue. This privilege is already implicit in the Aristotelian distinction we just invoked, since it is the exemplar of kindness who defines how ‘the’ kind person performs kind acts (e.g., for the right reasons and wholeheartedly). In that sense, the exemplar of kindness *is* ‘the’ kind person.

I am assuming that exemplars of kindness can sometimes fail to perform kind acts in situations where kindness is required. In other words, exemplars of virtue can be (somewhat) imperfect and this imperfection is fully consistent with their status as exemplars. Just how much imperfection is consistent with that status is reflected in the location of the threshold for ‘exemplar of kindness’—reflected, that is, in how far below one hundred percent it lies on the reliability continuum. I take it that the threshold for exemplar status lies somewhere between ‘highly reliable’ and ‘very highly reliable,’ but will not attempt to specify it further.<sup>15</sup>

Let me emphasise, however, that the margin of unreliability consistent with being an exemplar of kindness, whatever it may be, remains a margin of *moral error*. For it is a trivial truth that greater moral perfection is always better, morally speaking. Thus, while an exemplar of kindness is ‘permitted,’ after a fashion, to be less than perfectly reliable in the performance of kind acts, this permission does nothing to erase the fact that a less than perfectly reliable exemplar still has room to improve, morally. To that extent, no exemplar of kindness is also the measure of kindness (or not the final measure, anyhow).\*

The trivial truth about moral perfection was phrased in terms of what is better ‘morally speaking.’ I phrased it that way deliberately, since it is *not* a

\* It is not clear that this position is open to the metaphysical agent-centred view of virtue; and that may be regarded as an objection to it.

trivial truth that greater moral perfection is simply ‘better,’ full stop—nor, equivalently, that greater moral perfection is always practically normative or ‘to be pursued.’ Indeed, some philosophers have argued that this second proposition is actually false.\* Naturally, that sort of general scepticism about moral perfection is itself a ground for lowering the threshold to qualify as an exemplar of virtue. However, building imperfection into the threshold does not commit us to any such scepticism, since scepticism is not the only available ground.

Alternative grounds for a submaximal threshold for exemplar status may be found in plausible answers to a range of quite different questions. They may be found, for example, in plausible answers to the educational question, at what point of remoteness does a target’s being ‘out of reach’ overly discourage people from progressing towards it? Or to the philosophical question, how much, if anything, does morality have to concede to reality in order to preserve its authority? Or to the substantive moral question, how much forgivingness (in various senses) should moral standards incorporate? I suspect that the case for accepting some imperfection in an exemplar of virtue is overdetermined.

Exemplars of virtue are, in the first instance, ideals. Flesh-and-blood human beings will approximate these ideals to varying degrees, ranging from zero on upwards. Few human beings, presumably, achieve the status ‘exemplar of kindness,’ with most people falling somewhere on the reliability continuum between zero and, say, midway between the minimal threshold and the threshold for exemplar status.† The same goes for any other virtue. Now even this mild splash of realism raises a question, of course, about what practical relevance an exemplar of virtue can have for the rest of us, who are very unlikely to become one. For the time being, let us accept the naïve answer that the rest of us should still try to *emulate* exemplars of virtue. Later we shall examine its merits and also explore some alternative answers (in chapters 5 and 12).

In certain important respects, the distinction between our two thresholds on the reliability continuum resembles two other distinctions often employed in the analysis of virtue, namely, Aristotle’s distinction between natural virtue and full virtue, and the distinction between (moral) learners and experts.<sup>16</sup> Above all, it seems fair to say both that an exemplar of kindness has the *full* virtue of kindness and that she is an *expert* in matters of kindness. Moreover,

\* Most famously, in contemporary discussion, Susan Wolf (1982).

† Unlike some operational definitions of ‘character trait’ in psychology, virtuous character traits (i.e., virtues) are *not defined*, either in philosophy or in common sense, in such a way as to limit their possession to a minority. In principle, everyone may be an exemplar of kindness.

like someone not far above the minimal threshold for kindness, a person who has the natural virtue of kindness is only very imperfectly reliable about doing the kind thing. ‘Natural’ virtue, however, carries the connotation of being present from birth, whereas (a behavioural disposition’s correspondence to) the minimal threshold has no such implication. The ‘learning’ label maps even less well onto our distinction. For agents in the vicinity of the minimal threshold for kindness may not be trying to learn anything about being more kind (and, a fortiori, they may not be succeeding at it either). By contrast, exemplars of kindness may well—and, in any case, often should—be trying to learn how to be yet more kind. For these reasons, I shall usually retain the language of ‘exemplars’ of virtue.

§4. So far, I have been using kindness as my example of a specific virtue, while simply taking it for granted that kindness is, in fact, a genuine virtue. But is it? What about chastity? Or wittiness? Or (physical) strength? Asking which specific character traits, if any, count as genuine virtues—and more importantly, *why* they count or not—engages a very substantial philosophical question, one that has a distinguished history. It also represents a final way to hear the question, what is a virtue?

The traditional answer to this version of the question is given by *eudaimonism*. For our purposes, eudaimonism may be understood as follows:

(E) Trait<sub>1</sub> [T<sub>1</sub>] is a virtue if and only if—and because—an agent’s having T<sub>1</sub> contributes to her eudaimonia.

My formulation raises two obvious questions, namely, what does ‘contributes’ mean and what does ‘eudaimonia’ mean? To begin with the latter, and cutting brusquely through certain controversies, let us translate ‘eudaimonia’ as *flourishing*. (Happiness is the standard alternative.) How to explain ‘contributes’ is more complicated, since the term is designed to cover a range of variations within the theory. But to take the most straightforward case as an illustration, it means that an agent *has T<sub>1</sub> if and only if the agent is eudaimon* (i.e., flourishes),<sup>17</sup> so that we can rewrite (E) as

(E\*) T<sub>1</sub> is a virtue if and only if—and because—an agent has T<sub>1</sub> if and only if she flourishes.

For example, consider Rosalind Hursthouse (1999), one of eudaimonism’s clearest contemporary exponents. Hursthouse subscribes to (E), and comes close to (E\*), when she declares that ‘a virtue is a character trait a human being needs for eudaimonia, to flourish or live well’ (1999: 167).

I believe that (E) is fundamentally confused. Unfortunately, I do not have anything very satisfactory to offer as an alternative. As a result, I am forced to

fall back on the default answer, which may be fairly described as an intuitive mess. But let me at least explain why I reject (E).<sup>18</sup>

The Achilles's heel of eudaimonism is that it runs two questions together, thereby exposing itself to a special and extra-devastating version of what I shall call the 'pseudo-Prichard dilemma.'<sup>\*</sup> Specifically, (E) runs the headline question—How is the palm of virtue awarded?—together with another perennially vexing question: Does virtue pay?<sup>†</sup> These two questions do not have to be answered together. However, separating them is not an option for (E), since their conjunction is precisely what defines (E) as a distinctive strategy. (E) fixes the answer to 'Does virtue pay?' at 'yes,' and then awards the palm of virtue to all and only those traits that 'pay' in the relevant sense.

'Prichard' famously argues that attempts to answer the question Why be moral? all bottom out in an appeal to one of two things. *Either* they appeal to the agent's nonmoral self-interest (i.e., his happiness) *or* they appeal to the idea that virtue is its own reward (i.e., virtuous acts should be performed for their own sake). By arguing that each of these appeals is subject to a decisive objection, 'Prichard' introduces a dilemma. For convenience, mildly garnished with self-interest, let me label its two horns the 'signature horn' and the 'trailing horn,' respectively.<sup>19</sup>

On the *signature horn* of 'Prichard's' dilemma, we actually find a pair of objections. His first objection is that appeals to an agent's happiness (to explain why the agent should be moral) require an invariable coincidence between morality (or virtue) and the agent's nonmoral self-interest, but this coincidence fails to obtain. To this standard objection, 'Prichard' adds the distinctive objection that, even if the required coincidence does obtain, the self-interested reason the agent thereby acquires to perform the relevant acts—his incentive, if you like—is not really a reason to *be moral* (e.g., not a reason to recognise any *obligation* to perform certain acts). Appeals to the agent's nonmoral self-interest therefore fail to answer the question, properly considered.

In one sense, then, the 'proper' answer to the question is that 'virtue is its own reward.' The trouble, alas, is that some people find this answer unconvincing (many people?). It leaves them cold or, at least, insufficiently moved. What these people really want to know is, What *other* reason is there to be moral?

<sup>\*</sup> I build 'pseudo' into its title because my analysis of the dilemma is not entirely faithful to Prichard's (1912) text. (There are some details in endnote 19.) To suit our present purposes, I also frame the analysis largely in terms of virtue, specifically, rather than morality generally (as Prichard does). In any case, the main point(s) of interest are independent of the historical details.

<sup>†</sup> Or, to give the second question a possibly more familiar cast, why be moral?

Moreover, it is this second version of the question that is largely responsible for the urgency or resonance of the original. So, on the *trailing horn* of ‘Prichard’s dilemma, the objection is that the question *Why be moral?* already presupposes in effect that appeals to virtue being its own reward will fail to persuade.

Accordingly, if ‘Prichard’ is right, attempts to answer ‘*Why be moral?*’ must either fail to answer the question properly or fail to persuade (everyone). He concludes that we should reject the question instead. Since ‘Prichard’ also regards moral philosophy as being characteristically dedicated to this very question, he arrives at his eponymous conclusion that moral philosophy rests on a mistake. Thus, run-of-the-mill pseudo-Prichard. Let us now return to (E).

To begin with, we need to register a further philosophical issue about interpreting the concept of eudaimonia (however one translates the word). The issue is whether there is any restriction on how its ultimate constituents may be defined—in particular, whether *moral* constituents are eligible. Can having the virtues, for example, count as at least one ultimate constituent of the (most) flourishing life?<sup>20</sup> In the abstract, this is an open question, which may be debated on the merits. But in the specific context of (E), the question is closed. As far as (E) is concerned, none of eudaimonia’s ultimate constituents can be moral.<sup>21</sup> For moralising the (operative) definition of eudaimonia would trivialise the relation between virtue and eudaimonia, with the result that contributions to an agent’s eudaimonia would lose any traction they may have offered for explaining which traits are virtues. While this is manifest in the case where virtues themselves are among the ultimate constituents of eudaimonia—since then one would already have to know whether a given trait was a virtue in order to evaluate whether having that trait ‘contributed’ to the agent’s eudaimonia—it holds equally with other kinds of moral constituents.<sup>22</sup>

Hence, simply to get off the ground, (E) has to restrict its operative concept of eudaimonia to *nonmoral* constituents. It follows that (E)’s answer to ‘Does virtue pay?’ is forced onto the signature horn of pseudo-Prichard’s dilemma.\* On pain of equivocation, the currency in which (E) affirms that virtue ‘pays’ has to be the same as the currency in which the fact that a trait ‘pays’ is supposed to qualify the trait as a virtue. Since the latter currency is nonmoral eudaimonia, what (E) affirms, more fully, is that virtue pays in the currency of nonmoral eudaimonia. But this answer is clearly subject to both of the objections ‘Prichard’ deploys on his signature horn.

Worse, (E)’s criterion for awarding the palm of virtue inherits the same objections. We can take them in reverse order. Even if kindness, say,

\* Eudaimonism’s vulnerability to this dilemma is thus ‘special’ insofar as (E) is impaled on its signature horn *whether or not* there is a good objection lurking on its trailing horn.

invariably coincides with an agent's nonmoral eudaimonia, the most this establishes is that she has good reason to acquire that trait. It does not establish that kindness is a *virtue*, i.e., that the trait is any part of morality. Moreover, the claim that kindness (or any other virtue) invariably coincides with an agent's nonmoral eudaimonia is, of course, seriously implausible anyhow. The twists and turns in the formulation of eudaimonism—the various attempts to improve on (E\*)'s construal of the 'contributes' in (E)—mainly aim to cope with this problem (see, e.g., Hursthouse 1999, chap. 8).

A final, extra-devastating objection targets (E)'s distinctive move of awarding the palm of virtue to all and only those traits that pay in nonmoral terms, *under the presupposition* that virtue pays in nonmoral terms. Let us assume, for argument's sake, that justice is a genuine virtue. Now consider apparent counter-examples to (E\*)'s entailed claim that an agent's nonmoral eudaimonia invariably coincides with justice. Some will be cases like those raised by Hobbes's Foole, say, in which occasional promise breaking (injustice) at least appears to maximise an agent's nonmoral eudaimonia. For present purposes, it does not matter whether the counter-examples ultimately succeed against (E\*)'s entailed claim or not. What matters is rather our intuitive reaction to the apparent non-coincidence, namely, 'aha, so virtue does not always pay (not this virtue, anyhow).' Our reaction, in other words, is *not*: 'aha, so justice [strict promise-keeping] is not a virtue.' Still less is it, 'aha, so mild injustice [occasional promise breaking] is a virtue.' However, if (E\*) were the correct test of virtue, both of these other reactions would be fully warranted by the appearances in the Foole's case.

It is worth generalising this point a little: Faced with an apparent counter-example to (E\*)'s explanation of justice's hypothesised status as a virtue, our immediate reaction is clearly to fault (E\*)'s presupposition—and thereby, its criterion of virtue—as opposed to rescinding the assumption that justice is a virtue. That is to say, when push so much as threatens to come to shove, we baulk, happily and evidently, at (E\*)'s forced marriage of answers to 'How is the palm of virtue awarded?' and 'Does virtue pay?' Yet (E\*) shares this presupposition with (E); and (E) is equally nothing without it.<sup>23</sup>

§5. On this basis, I reject eudaimonism as an account of which specific character traits count as genuine virtues. I wish that I had a decent theory to offer in its place. But, alas, I do not. Accordingly, in discussing specific virtues, I shall restrict myself to conventionally plausible examples, and leave the deeper philosophical questions about their credentials as virtues unanswered.<sup>24</sup>

We should notice, however, that the reliance on intuition dictated by this theoretical modesty (or barrenness, if you prefer) occurs at the level of individual acts of kindness—specifically, at the level of paradigmatic or stereo-

typical acts of kindness—rather than at the level of any character trait directly. I do not mean to suggest that this makes the reliance on intuition any better per se. But it does mitigate some other concerns, about the value of a virtuous character trait's being instrumental. Let me explain.

I have assumed the metaphysical act-centred view, on which kind acts are basic. As I understand it, this means that the original locus of the moral goodness (or value) in kindness lies in individual kind acts (rather than in the trait of kindness). Kind acts, that is, have this value independently of anyone's disposition to perform them. However, on this analysis, the trait of kindness still acquires an instrumental value,\* since it reliably issues in independently valuable acts. Some philosophers object to qualifying traits as virtues merely on the basis of (their) instrumental value (e.g., Adams 2006, chap. 4). In their view, only an intrinsic value can qualify a trait as a virtue, and they object to standard forms of virtue consequentialism, for example, for flouting this requirement. One might therefore be concerned that the view I have described is open to the same objection, since that view has it structurally in common with virtue consequentialism that virtuous traits inherit their value from the acts they reliably produce.

Still, an important difference on this specific point remains. With virtue consequentialism, the instrumental value that a virtuous trait inherits from the acts it reliably produces is the very feature that qualifies the trait as a virtue.† On my view, by contrast, while a virtuous trait inherits both its value and its status as virtuous from the acts it reliably produces, each of these features is inherited separately. In other words, the former is not the basis of the latter. Rather, when a trait is virtuous, it inherits this status directly from the acts it reliably produces, i.e., from the fact that they already have that status themselves. Hence, on my view, it actually matters that the acts a trait like kindness reliably produces are specifically *virtuous* acts, rather than any old valuable acts.‡ It is the fact that the relevant individual acts—e.g., of helping an old lady to cross the street or of overlooking some fault in another—already have the status of virtuous acts that I am relying on intuition to 'secure.' Or again, if you prefer, that is what I am simply asserting, without explaining. However unsatisfactory that may be, it should at least be clear that it does not flout any requirement restricting the qualifications for virtue to intrinsic value(s).

\* For ease of exposition, I am writing as if this is the only value that the trait has. This assumption can be discarded and will be soon enough.

† Moreover, this value is really inherited from the *outcomes* that these acts produce in turn. For virtue consequentialism, the acts themselves are just (or at least, primarily) another intervening link in the chain.

‡ A fortiori, they must also be *acts*, and not merely conduits for outcomes.

Of course, if there is some such requirement, it presumably applies to individual virtuous acts as well as to traits. Thus, classifying ‘helping an old lady to cross the street,’ say, as a virtuous act would presuppose that this act was intrinsically valuable (or that some aspect of it was). Now the standard way to argue that such a presupposition is satisfied would be to claim that an agent’s intention to help an old lady for her own sake is intrinsically good or that his motive in helping her is (see, e.g., Hurka 2001, chap. 1); and no doubt they are intrinsically good. I have not affirmed separate good intention or good motive conditions on individual virtuous actions, but only because they seem unnecessary. Earlier I argued that, at some point on the reliability continuum, the goodness of a reliably kind agent’s motive in acting kindly emerges as a side effect of her reliability.\* There is a point on that continuum, that is to say, above which agents cannot ascend unless their motives in acting kindly are good ones. If I am right about that, then above this same point a reliably kind agent’s kind actions will be well motivated anyhow, and so intrinsically valuable (in conformity with the requirement in question).

Adams (2006) and Hurka (2001) both argue, though each for his own reasons, that virtuous traits have their own *intrinsic* value, in addition to their instrumental value.<sup>25</sup> I certainly accept that virtuous traits also have intrinsic value,<sup>26</sup> although it seems to me that the value of the acts a virtuous trait produces remains more important than its own intrinsic value, making its overall value largely instrumental. Adams and Hurka rightly emphasise, furthermore, that the intrinsic value of a virtuous trait often finds expression outside the context of virtuous *action*. This has been a leitmotiv of virtue theory reaching all the way back to Aristotle. I have paid little attention to this dimension of virtue, but that is not because I deny or denigrate its significance.† It is simply an artefact of having elected to concentrate squarely on the central relation between virtuous traits and virtuous actions.

§6. Let me return, finally, to the second half of my epigraph, *credo virtutem nec omnipotentam*. In many respects, the rise to philosophical prominence that virtue ethics has enjoyed over the last several decades has been a salutary development.‡ Nevertheless, it has had at least one unfortunate consequence, namely, that any philosopher who writes about virtue nowadays is liable to be read as a participant in the virtue ethics enterprise. More specifically, he or she

\* I focus here on the stronger of the two redundancy arguments from §2.

† On the contrary, as should become apparent, the view I defend accommodates this dimension very naturally and easily. Recall my title.

‡ Crisp and Slote (1997) offer a canonical collection of influential papers.

is liable to be read as subscribing to the imperialist ambition of that enterprise.<sup>27</sup> By this I mean its ambition to provide a complete moral theory, or a complete account of morality, centred fundamentally on the concept of virtue. Once again, the clearest illustration of this aspect of virtue ethics is provided by Rosalind Hursthouse, who advances the following formula: 'An action is right iff it is what a virtuous agent would characteristically (i.e., acting in character) do in the circumstances' (1999: 28).

I have no imperial ambitions. That is what I mean by saying, 'virtue is not almighty.' In particular, I believe that the virtues are only one province of morality among others. They are important and interesting, but not more so than some other provinces of morality, let alone all others. Most obviously, for example, *rights* are another province of morality; and rights are distinct from virtues. Moreover, I do not believe that the moral substance of rights can be reproduced or reconstituted in the language of virtue either, at least not more fundamentally.\* I shall not attempt to argue for this proposition, which I am affirming here primarily in the spirit of clarity in advertising: the account of virtue I offer is not intended to enable anyone to dispense with appeals to rights. (I discuss this a little further in chapter 4.)

Nor is my aim to describe some distinctive approach to ethics, to rival consequentialism or deontology or some other view.<sup>28</sup> My aim is rather to describe some truths about virtue, which are also therefore truths about morality. The extent to which these truths can be embraced or accommodated by established and well-known moral theories is mainly a matter of how flexible other theories are or how imaginative and open-minded their adherents are. I shall not enquire into how far that may be the case. But if some other theory and its adherents can accommodate all of what I claim, more power to them.

\* As it happens, I also believe that this truth is symmetrical: the moral substance of the virtues cannot be reproduced or reconstituted in the language of rights (or of other moral concepts), at least not entirely. I discuss this a little further in chapter 11.

## INDEX

- act-centred (priority) view, 3–5, 19, 90, 106, 243–65, 275, 331n2, 363nn1–3; defined, 3
- Adams, R., ix, 19–20, 80, 209, 331n9, 333n18, 334nn25–26, 339n3, 359n20
- adverbial requirement(s) on virtue, 6–7, 124–28, 196, 198–99, 334n2, 346nn18, 20, 360n26
- advice, 168–70, 180–85, 193–97, 237–40, 259–64, 301–4, 363n49, 364n13
- affect programs, 54–59, 63–66, 153–54, 223–24, 337nn23–24, 347n5
- agent-centred (priority) view, 3–5, 174, 244, 254, 331nn2, 5, 363nn1–2
- all things considered (ATC) permissibility, 8–11, 24–25, 92–94, 121–25, 138, 217, 235–40, 256, 263, 286–94, 315–25, 332n12
- Anscombe, G. E. M., 80, 91
- appropriateness: emotional vs moral, 61–62, 184–85, 187, 191, 194–95, 227–29, 264, 345n14, 347nn3, 8, 353n20, 354n28
- Aristotle, xii, 2–3, 5–6, 12, 14, 20, 23, 32, 34, 68–70, 72, 81, 121, 124, 126–27, 143, 149–51, 201–6, 212, 215, 220, 231–32, 235, 300–301, 331nn2–3, 5, 332nn15–16, 333n20, 335n11, 340n14, 355nn1, 3–4, 7, 356n10, 357n12, 362nn40–41, 363n2, 368n30, 370n10
- attention, 39–40, 46, 132–33, 152, 192, 253, 349n18, 354n25. *See also* emotional salience
- Averill, J., 47
- Badhwar, N., 67
- basic emotions, 41, 53–59, 63–66, 200, 248, 337n21
- Batson, D., 99, 112–20, 135, 137–38, 143, 158–59, 187, 247, 344n3, 345nn7, 9, 11, 347n4, 350nn23–25, 371n13
- bento box view (of moral psychology of virtue), 26–30, 284–85, 335n7; of compassion, 26–30, 125–28, 129–31, 138, 143, 165–67, 306, 351n27; of courage, 218–19, 222, 227
- bias. *See* out-group bias; racial bias in empathic neural responses
- black box view (of moral psychology of virtue), 26–30, 32–33, 125–28, 129–31, 196–99, 214–15, 231–32, 264–65, 334n4, 335n7, 359n22, 360n26, 362n39, 365n20
- blackspots (in virtue), 102–5
- blindspots (in virtue), 101–4
- Blum, L., 123, 136–37, 344nn4, 6, 345n10, 346n1
- Bommarito, N., 335n12
- Brady, M., 46–47, 188–92, 228, 251, 335nn1, 3, 336n12, 353nn22–25, 354n25
- Burnyeat, M., 68, 301, 303, 370n10
- Cao, Y., 162
- central test of virtue (CTV), 22, 25–33, 105–6, 334n2, 335n7, 349n18; defined, 24; for compassion, 125–26, 129–31, 135, 138–40, 146–48, 150–51, 162–67, 196–98, 265, 349n14, 351n29; for courage, 219, 214–23, 231
- character education: xiii, 150, 275–77, 279–80, 285, 297, 299–326; Fosbury model of, 304–6 (defined), 307–12, 314, 319–20, 370n7, 372nn21–22, 26, 373n31; ham-fisted Aristotelian model of, 300–301 (defined), 302–4, 306–12, 314, 371nn15, 18, 372n23

- cleverness, 28, 30–31, 70–73, 76–77, 129–31,  
143–44, 147–53, 162–67, 181–86, 194–97,  
218–23, 227, 254–58, 285, 326, 348n10,  
349n15, 351n28, 354n26; defined, 70
- coherentism, 295–97
- compassion (“v” subscript explained), 122
- compassion fatigue, 305–6, 371nn13–14
- conflict of desire, 6, 201–7, 212–13, 220–21,  
232–34, 355n3, 356nn9–10, 362nn40, 43–44.  
*See also* moral conflict
- continence. *See* strength of will
- Contreras-Huerta, L., 160–61
- Crisp, R., 20, 91, 334n28, 352n14, 368n30
- cross-situational consistency, 84–88, 95–107,  
313, 341n3, 343nn21–22, 344n24
- Damasio, A., 38
- Dancy, J., 366n12, 369n41
- Darley, J., 99
- D’Arms, J., 61, 336n11, 345n14, 347n8
- Darwin, C., 55–56, 64
- Decety, J., 305, 371nn13–14
- deference, 168–80, 192–96, 237, 261–64,  
352nn6, 9, 11, 15, 353n24, 365nn18–19;  
defined, 170
- degrees of reliability. *See* reliability continuum
- Deigh, J., 41, 336nn6, 10, 345n16
- Dennett, D., 38; his robot tale, 38, 46, 59, 134,  
136, 253
- Dent, N., 126–27, 214, 235, 335n10, 357n14
- Derryberry, D., 63, 338n32, 346n2
- de Sousa, R., 37–38, 46–47, 59–60, 153, 253,  
336n6
- dilemmas, 74–75, 340nn13, 15, 373n29.  
*See also* moral conflict
- direction of priority, 1–5, 175, 243–65, 331n3
- disunity of virtue, xi, 67–83, 97, 341n4;  
moderate, xiii, 79–80, 141, 144 (defined),  
163, 184, 193–94, 217–19, 236, 264, 316,  
321–26, 348n14, 351n1, 353n18, 373n31; radical,  
xiii, 78, 141, 144 (defined), 217–18, 235, 249,  
316–21, 325
- division of moral authority, 76, 145–49, 155,  
213, 218, 316, 318
- Doris, J., 87–89, 100, 259, 313, 341nn1–2, 4,  
342nn6, 10, 343n14, 364n12
- Dovidio, J., 156–59, 161, 348n10, 350nn20–23, 25
- Driver, J., 331n9, 332n10, 334n27, 352nn6, 13
- Dworkin, R., 296
- education. *See* character education
- Eisenberg, N., 120, 248, 344n2, 347n4
- Ekman, P., 46, 53–60, 63–66, 223, 248, 335n2,  
337n24, 338nn25–27, 30, 33, 34
- emotional salience, 37–38; input salience,  
46–50, 52, 57, 59, 133–36, 151–55, 183, 191–92,  
201, 223–25, 361n32; output salience,  
59–66, 133–34, 183, 191–92
- empathy. *See* sympathy
- empathy-altruism hypothesis, 112, 137, 187,  
247, 344n1
- empirical modesty, 88–95, 313, 342nn6–7,  
364n12
- Enoch, D., 171, 352nn6, 12, 15
- eudaimonism, xii, 15–18, 236–37, 332n17,  
333nn18–24, 347n7
- executive virtues, 70–71, 142, 200, 210, 212–14,  
216–18, 228, 231–32, 234–35, 316, 357nn14–15,  
358n17, 359n18, 367n20; defined, 209
- expertise, 14–15, 172–76, 193–97, 237–40,  
261–64, 332n16, 352nn7, 10–11, 362n47,  
363nn49, 52, 364n16
- externalism (about virtue), 8–12, 331n9,  
332n10–11
- facial expression of emotion, 39, 338n34;  
pan-cultural, 53–59, 64–66, 248,  
338nn25, 27
- Faucher, L., 46, 48, 54, 337nn16, 23, 338n24
- Flanagan, O., 82
- Foot, P., 78, 80, 355n4
- Frijda, N., 42, 47, 60–61, 338n30
- fundamental attribution error, 101, 107,  
344n25, 363n51
- Gaertner, S., 156–59, 161, 350nn20–22, 25
- Gelinas, L., 333n18
- Gleichgerrcht, E., 305, 371nn13–14

- Goldie, P., 39–40, 113, 116, 119, 184, 337n20, 338n31, 345n8, 346nn19–20, 353n20
- Greenspan, P., 336n8
- Griffiths, P., 54–55, 63, 336nn9–10, 337n24, 338n24
- Habermas, J., 367n26
- Han, S., 160–62, 350n26, 361n35
- Harman, G., 92, 341n2
- Hartshorne, H., 85–8, 95–100, 105, 107, 342n5, 343n17
- Heider, K., 59, 224
- Heyd, D., 291, 368nn28, 31, 34, 369n37
- Hills, A., 171, 177–78, 351n4, 352n13, 353n24
- Hodges, S. D., 306
- Hoffman, M., 115, 345n12, 347n4
- Hurka, T., 1–3, 20, 91, 245, 331nn1, 3, 333n18, 334nn25–27, 342n8, 343n14, 363nn1, 5
- Hursthouse, R., 15, 18, 21, 126, 273, 333n23, 334n28, 339n8, 366nn11, 13
- ideal(s), xi, 6, 14, 88, 177–80, 232–24, 260, 352n16, 363n4
- imperialism (about virtue), x–xii, 20–21, 334nn27–28
- integral view of virtue (IV), xi, 30–31, 33, 111, 282, 284; for compassion, 30–31, 111, 123–31, 193–98, 264; for courage, 30, 199–200, 214–31
- Irwin, T., 5, 67, 126, 339nn1–2
- Isen, A., 99
- Jacobson, D., 61, 336nn1, 345n14, 347n8
- James, W., 41–42, 336n5
- justification. *See* moral justification
- Kamm, F. M., 292–94, 368n33, 369nn36–38
- Kristjánsson, K., 32, 335n11
- Lange, C., 41–42, 336n5
- Lapsley, D., 370n6
- Latané, B., 99
- Lazarus, R., 50–52, 337nn18–19
- Lear, J., 207–8, 210, 356n11, 357nn11, 16, 358nn17–18, 359n18
- LeDoux, J., 48, 50–51, 200–201, 223–26, 337n18
- Levenson, R., 56–57, 338nn25–27
- Levin, P., 99
- MacIntyre, A., xiii, 281, 365n4
- Malhotra, D., 159, 350nn24–25
- margin of unreliability, 9–13, 332n12, 365n20
- Martin, R., 138
- May, M., 85–88, 95–100, 105, 107, 342n5, 343n17
- McDowell, J., xiii, 7, 69–70, 80, 95, 121, 249, 281, 283, 331nn5, 7–8, 363n2
- McGinley, P., 269
- Milgram, S., 99–104, 343n21
- Mineka, S., 223, 226, 361n33, 362nn37–38
- minimal moral decency, 79, 97, 103–5, 144–47, 151, 163–64, 166, 184–85, 218–19, 222, 227–28, 236, 240, 264, 286, 316, 321–26, 334n6, 339n10, 340nn11–13, 348n14, 349n16, 351nn27, 1, 353n18, 354n26, 363n52, 364nn8, 16, 365n19; defined, 80
- modest agent-centred (priority) view, 93–95, 106, 176, 243–65, 268, 270–71, 273, 275, 282, 291, 296–97, 301–2, 315, 343n15, 363n3, 366n9; defined, 4–5
- Mogg, K., 48, 63, 201
- moral conflict, 7, 25, 74–75, 96, 262–64, 286–94, 315–25, 339n5, 364n15, 369nn36–39, 370n18, 372n25, 373nn30–31
- moral courage, 224–27, 361n33
- moral justification, xiii, 266–98, 365n1, 367nn19, 21
- moral mistake, 68–77, 80, 121, 211–12, 214, 322–25, 339n3, 350n27, 359n19; internal and external, 141 (defined), 142–47, 163–65, 213, 217–18, 235, 349n15, 351n28
- moral principles, 266–98, 365n3, 367n23
- moral rectification (of emotion), 30–31, 123, 130, 148–51, 155–62, 223–27, 285, 361n36
- motives, xi, 1–3, 10–13, 20, 33–34, 91, 112, 117–18, 245, 247, 271, 301, 304, 331nn1, 9, 332nn13–14, 334n26, 335n13, 363n6, 370nn4, 10; emotional, 42–43, 60–62, 137, 149, 187, 219, 221–22, 232–33, 334n1, 345nn9, 11
- Murdoch, I., 133, 183, 346n1

- natural virtue, 14–15, 68–70. *See also* proto virtue
- Nichols, S., 112–16, 118, 120, 344nn2, 4, 6, 345nn7, 11
- Nisbett, R., 86–87, 96, 101, 344n25
- Nussbaum, M., 42–43, 122–23, 163, 255, 261, 336n7, 339n6, 345nn15–16, 346n1, 365n17
- Öhman, A., 48–52, 57, 155, 223, 225–27, 331nn17–18, 361n33, 362nn37–38
- Olsson, A., 226
- O’Neill, O., 4, 283–85, 366n4
- out-group bias, 145–47, 155–62, 164–65, 218, 348n12, 350nn20, 23
- paradigmatic acts (of virtue), 4–5, 7, 18–19, 24–25, 93–94, 98, 106, 132, 174–76, 207–8, 244, 254–57, 260–64, 275–77, 279, 296–97, 308–9, 315, 344n24, 346n17, 349n14, 363n3
- particularism, xiii, 366n12, 369n41
- Pears, D., 201–2, 206, 212, 215, 220, 232, 355n1, 356nn8, 10, 357n14, 362n43
- Pears-Aristotle prototype (PAP), 207, 212–13, 215, 220, 356n10, 359n21; defined, 206; half-PAP, 212 (defined), 213, 215–16, 220, 222
- perceptual theory (of emotion), 37, 41, 44–45, 188–91, 335n3, 336n12
- perfection, 13–14, 78, 100, 318, 321, 326, 366n9, 367n18
- permissibility. *See* all things considered (ATC) permissibility
- Phelps, E., 50, 225–26
- plasticity of emotion, 45, 53–59, 153–62, 337n23, 338n28; fixed-culture, 57–58 (defined), 154–55, 161, 224–27, 279–80, 361n36; variable-culture, 57–59 (defined), 63–66, 154–55, 161–62, 224, 338n29, 361n36
- Plato, xiii, 299
- Plenty Coups, 207–8, 210–11, 358n18
- Prichard, H. A., 16–17, 333n19
- Prinz, J., 43, 61, 336nn3, 13, 337nn18–19, 22
- priority. *See* direction of priority
- promises, 82–83, 235–36, 292–94, 369n36
- proto-authority, 168, 180, 194–97, 227–28, 264, 354n28; defined, 186
- proto virtue, 69–74, 77–78, 121–23, 141, 200–201, 339nn4, 7; defined, 69
- racial bias in empathic neural responses, 159–62, 361n35
- rationality of fear, 42–43, 201, 204–7, 212–13, 219–22, 231–34, 335n7, 356n9, 359n21, 361n30, 362n43; three grades of, 203 (defined), 220
- Rawls, J., 295–97, 369n41
- reciprocity of virtue, xi–xii, 5, 9–11, 22, 67–83, 121, 139, 141, 144–45, 184, 213, 217, 235–40, 263, 319, 321, 323, 325, 332n17, 334n6, 339nn2, 4, 7, 340n15, 348nn11, 14, 353n18; defined, 67–68
- reliability continuum, 12–15, 20, 34, 102–4, 106–7, 261, 300–312, 317
- right reasons requirement, 5–12, 95, 117, 197–98, 229–31, 331n6, 343n23, 354n29, 362n39
- rights, 21, 76–83, 286–94
- Roberts, R., 32, 209, 335n11
- Robinson, J., 41
- Ross, L., 86–87, 96, 101, 344n25
- Ross, W. D., 91, 334n25
- Rothbart, M., 338n32, 346n2, 347n2
- salience. *See* emotional salience
- salience problem, 133, 153, 218; for anger, 134–35; for fear, 219–23; for sympathy, 135–40, 162–64, 347n8
- Scanlon, T., 282–85, 295, 333n19, 367n23
- Scarre, G., 200, 208, 224, 355nn1, 4, 362n41
- Scherer, K., 47, 59, 63, 224
- Shade, P., 214, 235, 357n14
- Sheng, F., 160–61
- Sibicky, M., 348n10
- situationism, xii–xiii, 84–107, 258–61, 341nn1–3, 342nn6–7, 9, 343n17, 363n51, 364n12; rearguard of, 89–95, 313–15, 371n21, 372nn22–23
- Sliwa, P., 351n4, 352n6
- Slote, M., 3, 20, 363n1

- Sober, E., 114, 117  
Stark, S., 355n6  
strength of will, xii, 6–7, 23–24, 149–51,  
201–3, 206, 231–34, 334n2, 356n10, 360n26  
supererogation, 288–94, 368nn28–35,  
369nn36–38  
supplementary moral knowledge, 28, 30–31,  
129–31, 147–53, 162–67, 218–19, 222, 227,  
255, 285, 321, 326, 334n6, 349nn27–28,  
354n26, 360n27, 364n16, 365n19  
sympathy: distinguished from compassion,  
27–28, 112, 120–24, 138–40, 142–43, 184,  
247–50, 255–56, 347n6, 365n17; distin-  
guished from empathy, 112–20, 344nn3–6  
  
tack-on view (of moral psychology of  
virtue), 126–29, 196, 346n19. *See also*  
black box view (of moral psychology of  
virtue)  
Tappolet, C., 37, 39–46, 48, 54, 60–62, 335nn1,  
3, 336nn4, 6–8, 12, 337nn16, 23, 338n24,  
355n2  
teacher (of virtue), 301–4, 309, 370n7  
temporal stability, 84–87, 100–102, 105,  
341n3, 343nn21–22  
thick deontic description, 92–95, 169, 182,  
244–45, 269, 271, 286, 290, 364n15;  
defined, 91  
Thomson, J., 3, 76, 90–93, 331n3, 343n14,  
363nn1, 5, 367n25  
threshold of emotion trait, 134–35, 346n2;  
for fear, 219–20, 223, 227; for sympathy,  
137–38, 140, 143, 146, 166–67, 306  
  
threshold(s) of virtue, 12–15, 20, 34, 102–4,  
106–7, 259–60, 300–312, 317–19, 365n20,  
370n15  
Toi, M., 115, 345n7  
  
understanding, 170–71, 175, 177–80, 352n13,  
370n8  
unity of virtue. *See* reciprocity of virtue  
Urmson, J. O., 368nn34–35  
  
virtue consequentialism, 19–21, 331n9,  
332n10, 333n18, 334n24, 368n27  
virtue ethics, x, xii, 20–21, 75–76, 84, 258,  
260, 334nn27–28, 364n12  
virtue imperialism. *See* imperialism (about  
virtue)  
  
Wallace, J., 209, 213, 359n20  
Walzer, M., 281, 365n4  
Wang, C., 161  
Watson, G., 3, 340n12, 363n1  
weakness of will, 6, 22–24, 31, 121, 150–51,  
334n1, 367n24  
wholeheartedness, 5–7, 124, 127–28, 149–51,  
201, 232–34, 334n2, 356n10  
Wiggins, D., 132, 346n1  
Williams, B., 6, 81, 91, 331nn6, 8, 335n8, 365n2  
Wolf, S., 14, 67, 70, 339n2  
  
Xu, X., 160–62  
  
Zajonc, R., 50–52, 337nn18–19  
Zuo, X., 162, 350n26, 361n35