

Virtue and Reliability
Jason Baehr
Loyola Marymount University

Abstract: This paper examines whether reliability is an essential or defining feature of a (moral or intellectual) virtue. I defend four main theses. I argue (1) that reliability is *not* a defining feature of a virtue where virtues are conceived (as they often are) as “personal excellences,” but (2) that there is another (also intuitive and familiar) conception of a virtue according to which reliability *is* a defining feature. I also argue (3) that even on the former conception a certain rational *belief* pertaining to reliability is essential and (4) that reliability itself, while not a defining feature a virtue thus conceived, nevertheless is a concomitant of it.

My concern here is with what we might call the *logic* of a virtue, that is, with the *essential* or *defining* features¹ of a virtue. Therefore in contrast with, say, Aristotle’s inquiry into virtue, I am not necessarily concerned with qualities that virtuous people as we know them tend to possess, but rather with qualities that *any* virtuous person *must* possess. Specifically, I am concerned with whether “reliability” – or the reliable achievement of the ends or goals proper to a virtue – is an essential or defining feature of a virtue.

By a “virtue” I mean an excellence of personal *character*. And my interest lies with both moral and intellectual virtues. Moral virtues, as I am thinking of them, are excellences of moral character that include traits like courage, generosity, honesty, and compassion. Intellectual virtues are excellences of intellectual character and are essentially the intellectual counterpart to the moral virtues. They include traits like *intellectual* courage, caution, perseverance, carefulness, thoroughness, openness, and

¹ I say “essential” or “defining” features rather than *necessary* features because I will argue that reliability is in a certain sense a necessary feature of a virtue even though it is not part of what *makes* something a virtue.

fairness.² I am interested, then, in whether it is a defining feature of moral and intellectual virtues that their possessor reliably achieve the ends or goals proper to each kind of virtue.

There is considerable disagreement about this issue in both ethics and epistemology. One need not look far to find prominent figures in each field sitting on opposite sides of the fence. In ethics, for instance, Julia Driver³ claims that reliability is *the* defining feature of a moral virtue, while Michael Slote⁴ argues that reliability is not essential. In epistemology, Linda Zagzebski⁵ argues that reliability is a necessary condition for an intellectual virtue, while James Montmarquet⁶ denies this claim.

I defend four main theses in connection with this issue. I argue (1) that reliability is *not* a defining feature of a virtue where virtues are conceived (as they often are) as “personal excellences” or as traits the possession of which make one a good or better person, but (2) that there is another (also intuitive and familiar) conception of a virtue according to which reliability *is* a defining feature. I go on to argue (3) that even on the former conception a certain rational *belief* pertaining to reliability is essential and (4) that reliability itself, while not a defining feature of a virtue thus conceived, nevertheless is a concomitant of it. The discussion sheds important light on the nature of intellectual and moral virtues and on certain debates in epistemology and ethics.

I

² I am therefore thinking of intellectual virtues along “responsibilist” rather than “reliabilist” lines. Proponents of the latter conception tend to identify cognitive *faculties* and related abilities like vision, memory, introspection, and reason – rather than any *character* traits – as paradigm cases of intellectual virtue. See Axtell 1997 for more on this distinction.

³ Driver 1996. See also Jeremy Bentham’s *An Introduction to the Principles of Morals and Legislation* (1789).

⁴ Slote 1997 and 2001.

⁵ Zagzebski 1996.

⁶ Montmarquet 1993 and 2000. Jonathan Dancy (2000) defends this position as well.

Before getting to the main argument, I must draw attention to two additional features of moral and intellectual virtues as I shall be thinking of them. The first concerns the ends or goals proper to each. On the general conception of moral and intellectual virtues assumed here, an intellectually virtuous person is one who is motivated in a particular way by cognitive or intellectual ends like truth, understanding, and rationality. Such a person characteristically desires, aims at, cares about, takes pleasure in, etc.,⁷ these values. By contrast, a morally virtuous person is ultimately motivated by certain distinctively moral ends. He cares about, pursues, delights in, etc., things like the promotion of justice, human flourishing, and the alleviation of suffering.⁸ We need not worry here about trying to specify the complete set of values proper to each kind of virtue. The important point is that in asking about the reliability of these virtues we are asking about their reliability vis-à-vis values or goals of the sort just specified.

Second, I shall be assuming, at least to begin with, a general conception of these traits according to which they are “personal excellences,” which again means that one is a better person or is better qua person on account of possessing them. While I cannot explore this notion in any detail here, I assume it is a familiar and intuitive way of thinking about the nature of a virtue. It has much in common, for instance, with accounts of virtue according to which virtues are admirable personal traits⁹ or constituents of

⁷ Throughout the paper I shall be describing the relevant psychological dimension of both moral and intellectual virtues in this pluralistic and open-ended way. I do not want here to take a stand as to exactly what the relevant psychological requirements are, e.g., whether to possess a virtue a person must possess certain desires or intentions or pleasures or some combination thereof. I take it as given, however, that *some* such requirements exist. In employing this pluralistic characterization, then, I am simply making reference to those requirements, whatever exactly they turn out to be.

⁸ Zagzebski (1996) draws a similar distinction between moral and intellectual virtues. She also makes the further (and to my mind less plausible) claims that intellectual virtues are a subset of moral virtues and that normative epistemology is a branch of ethics.

⁹ See, e.g., Slote *op. cit.*

human flourishing.¹⁰ The basic idea is that while certain traits of character may make one better qua, say, athlete, musician, carpenter, or attorney others have a more direct bearing on one's personal worth or excellence. One obviously can be an excellent musician or carpenter, for example, without being a particularly good person; conversely, one can be deeply compassionate, generous, or honest and yet be a failure as a musician or carpenter. Virtues as I shall be thinking of them, then, are those traits of character the possession of which contribute to one's personal worth or make one good or better qua person.¹¹ As the example above illustrates, this conception of a virtue is clearly applicable to those traits that we typically regard as moral virtues. But it also applies to traits commonly regarded as intellectual virtues. A person who desires the truth for its own sake, who is committed to acquiring a genuine understanding of a range of important topics, and who, in pursuit of these ends, makes substantial sacrifices, remains open and honest to viewpoints that initially seem to conflict with his own, inquires in a patient, careful, and attentive way, clearly would seem a better person as a result. This is not to say that this person would be *morally* better (though on a broad enough conception of morality this might be true as well), but there is little reason to think that the domain of *personal* excellence or worth must be limited to or exhausted by that of *moral* excellence or worth.¹²

¹⁰ The accounts of Aristotle and neo-Aristotelians like Rosalind Hursthouse (1999) are probably examples; however, it is not entirely clear whether either thinks of a virtue *merely* as a constituent of human flourishing, since on supportable readings of each, they think of virtues as *both* constituents of and a *means* to human flourishing. I will have more to say about such "mixed" or "hybrid" conceptions of virtue below.

¹¹ This is not to deny that we can speak sensibly of the "virtues" (even *character* virtues) of an athlete, musician, carpenter, etc. In fact I defend this point below.

¹² My own account, which I cannot explain in any detail here, is that personal excellence or worth is a function of certain intentional psychological relations between an agent and certain valuable (moral and intellectual) ends. To be good qua person is, for example, to love, care about, etc., certain moral values like justice or the alleviation of suffering, but also certain intellectual or cognitive values like knowledge and understanding. This explains how personal worth or excellence can have both a moral and an intellectual dimension. For a similar account, see Hurka 2001.

II

Our concern, then, is with what I shall call the “reliability thesis.” This is the claim that reliability of the sort just noted *is* an essential or defining feature of moral and intellectual virtues.

At first glance, there would appear to be a good deal of intuitive support for the reliability thesis, for normally we would more than balk at calling a person genuinely intellectually virtuous if, despite his best efforts, he never actually reached the truth, acquired knowledge or understanding, etc. Similarly, we would scarcely think of a person as genuinely benevolent if, good intentions notwithstanding, she regularly failed to promote the well-being of other people. Such individuals would likely strike us as *incompetent* rather than virtuous. Indeed, it might be wondered whether there is anything more to an ascription of virtue than a judgment that the person in question has a certain competence or is reliable at producing certain good outcomes.

While I will return to some of these considerations below, it would be a mistake to give them too much initial weight in the present context, for they mainly concern what we would expect of virtuous people in *our* world or in worlds *like* ours. To be sure, a person who, under normal conditions, failed regularly to reach the truth or to affect the world in a morally good way could not be considered virtuous. It does not follow from this, however, that such reliability is an essential or defining feature of a virtue; for again, the latter are qualities that *any* virtuous person *must* have, and it is at least questionable whether under *nonstandard* conditions, a person who failed in the ways just mentioned would also fail to be virtuous.

Indeed my first argument against the reliability thesis exploits this possibility. The argument centers around the following two cases.

Case 1: Imagine an epistemic agent, Tori, who inhabits a world controlled by an epistemically malevolent Cartesian demon. Tori has all the “internal” marks of an intellectually virtuous person: she desires the truth, is committed to achieving an understanding of important issues, is willing to make sacrifices so that she can achieve her epistemic goals, is tenacious and patient in inquiry, listens fairly and openly to others, evaluates evidence carefully and thoroughly, etc. By her lights and by the lights of everyone around her, Tori is extremely intellectually successful: she has, let us say, enjoyed a long and influential career as a genetic researcher producing excellent work that has advanced her field in important ways; she also reads voraciously, has an impressive sense of local and global history, geography, and politics, a sophisticated grasp of various complex issues in economics, medicine, ethics, etc., and is uniquely adept at making insightful connections between these various bodies of (apparent) knowledge. Appearances aside, however, Tori is radically deceived. Given the real nature of things, which the Cartesian demon keeps perfectly hidden, most of her supposed “knowledge” is riddled with falsehoods.

From an intuitive standpoint,¹³ it seems fairly obvious, despite her overwhelming lack of cognitive success, that Tori is genuinely intellectually virtuous: that she is, for instance, genuinely intellectually tenacious, fair, patient, careful, reflective, etc. After all,

¹³ By an “intuitive standpoint” I mean simply a (more or less) commonsense or theory-neutral standpoint, and more specifically, the standpoint of one who is not already committed to a substantive account of whether reliability is a defining feature of a virtue. It is of course critical to any rational adjudication between competing answers to our central question that such a standpoint exists and that we are (more or less) able to adopt it.

she has all the right motives, intentions, feelings, etc., and does all she can to achieve her epistemic goals. Her *failure* to reach the truth is due entirely to her extraordinary and unfortunate circumstances – circumstances entirely beyond her control.¹⁴

Case 2: Now consider Tori's moral counterpart Ted, who is the victim of a *morally* wicked Cartesian demon. Upon returning from a recent, eye-opening trip to an impoverished third-world country, Ted has resigned as the CEO of a lucrative but soulless corporation to start a nonprofit organization aimed at improving the plight of various poor and oppressed people groups across the globe. Ted has done so, not out of guilt or any other self-oriented motivation, but because of a genuine, altruistic concern for the plight of the less fortunate, a concern which has come to occupy an entrenched and prominent place in his psychology. The demon, while not concealing the general nature of reality from Ted, nevertheless thwarts all of his moral efforts. Though Ted thinks that his fundraising is resulting in hundreds of thousands of dollars of aid being sent around the world, the demon systematically stymies the transactions, funneling the cash into slush funds at Ted's former corporation. While Ted continues to build the organization and to grow in his concern and compassion for the people whose lives he is working to improve, his efforts are almost entirely morally fruitless.

Again, insofar as Ted's moral outlook and sympathies are genuine, as long as he possesses the relevant motives and feelings and does what he can to achieve the relevant moral ends, it seems clear that he is genuinely morally virtuous. While his efforts

¹⁴ James Montmarquet (1993) and Casey Swank (2000) defend similar arguments against the reliability thesis.

ultimately are thwarted (and indeed may be doing more harm than good), this is in no way his own doing or fault, and so would not appear to make him less than genuinely compassionate, benevolent, generous, and so forth.¹⁵

While apparently providing good reason for thinking that reliability is not a defining feature of a virtue, these cases may carry little weight for someone already convinced of the reliability thesis. Such a person might claim, for instance, that because of the unfortunate *outcomes* of their efforts, Tori and Ted simply *cannot* be genuinely virtuous: Tori is not genuinely fair-minded or intellectually tenacious because she fails to consistently hit upon the truth and Ted is not genuinely compassionate and generous because he fails to actually improve the lives of those he is trying to help. While I will attempt later on to make some sense of this reply, for the moment I note simply that it comes at a high intuitive price. For again, from an intuitive standpoint, it is difficult to deny that Tori and Ted are genuinely virtuous even though they fail to accomplish their respective goals.

There is, however, an alternative move open to a defender of the reliability thesis, one that does not require a denial of the claim that Ted and Tori are genuinely virtuous. Note that Ted and Tori have character traits that in *other* possible worlds or environments *would* lead to very good epistemic and moral outcomes. Indeed, they fail to do so in reality only because Ted and Tori inhabit extremely bizarre and nonstandard worlds. We may distinguish, then, between two types of reliability: viz., “categorical” reliability and “indexical” reliability. A trait is *categorically* reliable just in case it is reliable in any possible world or environment, while a trait is *indexically* reliable just in case its reliability is indexed or relative to a limited set of worlds or environments. Accordingly, a

¹⁵ See Garcia 2003 for a similar line of reasoning.

defender of the reliability thesis could hold that moral and intellectual virtues are indexically (though not categorically) reliable and in doing so agree with the claim that Ted and Tori are virtuous. For again, they possess traits which in other possible worlds or environments would lead to certain good moral and intellectual outcomes.¹⁶

But there is a fairly obvious problem with this objection: namely, that from an intuitive standpoint, Ted and Tori are virtuous in light of who they are and what they do in the very worlds they occupy, that is, in light of who they are *considered in its own right*. It seems wrong to suggest that Ted and Tori are virtuous on account of what they might accomplish in certain other possible worlds (possible worlds very distant from their own). Rather, it is Tori's very love of knowledge and her willingness to do what it takes to acquire it – her love and willingness taken by themselves or considered in their own right – that make Tori genuinely intellectually tenacious, careful, reflective, etc. An analogous point holds for Ted: he is morally virtuous on account of his compassionate outlook and generous heart *themselves*, not because of what, in certain other, *nonactual* worlds, might result from these traits.¹⁷ Another way to see the point is to recall that we are thinking of intellectual virtues as personal excellences. It seems plainly false to think that Tori or Ted is a better *person* – especially a better person *here and now* – on account of certain results they *would* produce were they to find themselves in radically *different* environments. Such counterfactual considerations seem irrelevant to their present personal worth.

¹⁶ This kind of move is made by some virtue epistemologists in response to the so-called “new evil demon problem,” which concerns the epistemic status of the beliefs of a victim of a Cartesian demon. Intuitively, the person's beliefs are justified, which in turn seems to pose a problem for thinking that justification is a function of reliable traits or intellectual virtues. In response, Sosa (1991) and others claim that justification is a function of traits that are reliable merely in an indexical (rather than the categorical) sense just noted.

¹⁷ This is not to deny that it might also be reasonable to conclude that Ted and Tori are virtuous on the basis of their indexical reliability. Indeed, I defend such a point below. Rather, my claim here is that from the *intuitive* standpoint in question, the basis for this judgment is not the fact that they are indexically reliable but rather the very content of their psychology.

This leads to a second, more principled objection to the reliability thesis. Note again that the failure of Tori and Ted to accomplish their goals is *no fault of their own*; it is on account of terribly *bad luck* that their best efforts go awry. It would seem, however, that judgments about personal worth or excellence are largely *immune* to factors beyond an agent's control, that is, immune to the kind of epistemic and moral luck in question. It cannot reflect poorly on *Ted* (or on Ted as a *person*), say, that his good moral intentions and efforts are thwarted by a malevolent agent that he is entirely unaware of and can do nothing about. Ted might have a more desirable *life* were he not so unlucky (certainly others would), his *actions* might even be considerably more valuable (in fact, if act utilitarianism is true, this might be necessary for his actions to be considered morally *right*), but a morally friendlier world would not make Ted a better *person*. And of course an analogous point could be made regarding Tori: while she would enjoy greater intellectual success in an epistemically friendlier world, she would not thereby be an intellectually better person.¹⁸

If in fact personal worth cannot be determined by factors susceptible to the relevant kind of epistemic and moral luck, then it would seem that reliability cannot be a defining feature of a virtue conceived as a personal excellence. For, as the cases of Ted and Tori illustrate, our actual success in our intellectual and moral endeavors is often, in a very deep way, a matter of luck. Whether our inquiries lead to the truth frequently depends on a number of factors beyond our control: for instance, that other people are

¹⁸ As these comments suggest, I am denying the existence of only a certain, fairly narrow kind of moral and epistemic luck. I am claiming only that the extent to which one is good qua person in the relevant (moral or intellectual) sense is not a matter of a particular kind of luck. I do not deny that other kinds of normative status (e.g., the kinds noted above) can be a matter of luck. Nor do I deny that personal worth or excellence is a matter of luck *in certain respects*. It might depend, for instance, on the community in which one was raised or the kind of parenting one received. Thus I do not see any conflict between my argument and the claims regarding moral luck by authors like Bernard Williams (1981) and Thomas Nagel (1979).

telling us the truth, that “common knowledge” really can be trusted, that various received methods of acquiring information are reliable, and, of course, that appearances in general are not, as they are for Tori, radically misleading. Similarly for our moral efforts: success in this area requires, for instance, that the people we aim to help really do have the needs they appear to, that others bring to completion certain efforts that we initiate, and that our moral aims and efforts are not systematically thwarted in the way that Ted’s are.

Reliability, then, is a deeply chancy phenomenon. But since virtues as we are conceiving of them just *are* personal excellences, a person’s reliability or lack thereof cannot bear on whether she possesses a virtue, which is another way of saying that reliability is not an essential or defining feature of a virtue thus conceived.¹⁹

In response, a defender of the reliability thesis might take issue with the general conception of a virtue underlying this argument. It might be said that a virtue is not *identical* to a personal excellence, but rather that the notion of a virtue incorporates *both* the notion of personal excellence *and* the notion of reliability. By this account, which we might refer to as a “mixed” conception of a virtue, virtues have a kind of dual nature. To say that a person is virtuous is to say that she is both good qua person in the relevant sense and that she possesses a certain *competence*, that is, that she can be counted on to achieve or is *reliable* with respect to the ends or goals proper to the virtue in question. By holding to a mixed conception, a defender of the reliability thesis can agree that

¹⁹ This holds even where reliability is understood indexically. For even when intellectual or moral virtues are exercised in “suitable environments,” this typically is not entirely *sufficient* for achieving the relevant goals; other factors also play a role and at least some of these factors (e.g., the successful assistance of others) are beyond the control of the agent in the relevant sense.

reliability does not bear on personal worth in the manner indicated above while still maintaining that reliability is a defining feature of a virtue.²⁰

The first thing to note about a mixed conception of a virtue is that it attempts to bring together two very different sorts of values under a single, univocal conception of a virtue. Understood as personal excellences, virtues have their value “in themselves”; they are intrinsically valuable. The value of being good qua person is not, for instance, derivative from the value of the things such a person is likely to accomplish. Understood as competences or reliable dispositions, however, the value of virtues is instrumental; it is derivative from the value of the ends to which they are a reliable means. While this does not by itself constitute an objection to a mixed conception, it does raise the question of whether this conception is attempting to bring together what are actually two different *kinds* of virtue under a single, univocal conception of a virtue. I will return to this point below.

This initial worry aside, there is good reason to think that the requirements of a mixed conception are too strong. If reliability were in fact a defining feature of a virtue, then to assert that a person is virtuous would be (in part) to assert that he would, at least in certain environments, be successful at achieving certain valuable ends. But if this were the case, we could not reasonably make virtue ascriptions in the absence of a consideration of this kind of success or potential success. The problem is that we *can* be warranted in making such ascriptions absent the relevant kind of concern with or attention to the probable outcomes of a person’s character. As the cases of Tori and Ted indicate, it is possible to be justified in regarding a person as virtuous just on the basis of

²⁰ Mixed conceptions can be either strong or weak depending on whether they construe reliability categorically or indexically. Zagzebski (1996) defends a strong mixed conception with respect to both moral and intellectual virtues.

their virtuous motives, intentions, desires, etc., considered in their own right. It is not essential that we also consider the potential good that might result (and only in other possible worlds at that) from these qualities.

Similarly, if a mixed conception of a virtue were correct, knowing that a particular person possesses a character trait which under certain conditions would lead to the achievement of certain valuable ends also would be an insufficient basis, when taken by itself, for regarding that person as virtuous. But this too seems counterintuitive. It seems reasonable, that is, to think that if we know a person to have a trait of this sort, then quite apart from any consideration of the bearing the trait might have on this person's worth qua person, we shall be warranted in regarding this person as virtuous, at least in some legitimate sense of the term. After all, we can talk sensibly about the virtues of, say, an athlete or musician or carpenter, and yet none of these traits would appear to have any immediate bearing on the personal worth or excellence of their possessor. This is attributable to the fact that in the broadest sense a virtue is simply an *excellence*, and there is more than one way in which a character trait might be good or excellent. It might be good because it contributes to personal worth or excellence or it might be good (as in the case of the virtues of an athlete, musician, etc.) simply because of its favorable outcomes.

This suggests a distinction between what we might call a "personal worth" conception of a virtue, according to which a virtue just is a personal excellence, and a "competence" conception, according to which a virtue is simply a competence or disposition to bring about certain valuable ends in a reliable way. Indeed, while not a widely recognized distinction, there is further evidence of it in the literature of both virtue

epistemology and virtue ethics. Virtue epistemology, for instance, is routinely divided into two main camps.²¹ The primary basis for the division is the camps' differing conceptions of an intellectual virtue. "Virtue reliabilists" conceive of intellectual virtues (roughly) as any reliable quality of a person, while "virtue responsibilists" conceive of them as the traits of a responsible inquirer. While these groups regularly quibble over who has the "right" conception of an intellectual virtue,²² a more plausible diagnosis of the situation is that neither group has a monopoly on the concept, that each is dealing with its own viable and legitimate general conception of an intellectual virtue (i.e., that virtue reliabilists are occupied with a competence conception and virtue responsibilists with a personal worth conception).²³ A similar point can be made in connection with virtue ethics. As noted earlier, Julia Driver identifies reliability as the exclusive virtue-making property of a moral virtue, while Michael Slote analyzes this notion strictly in terms of certain admirable and intrinsically valuable motives. Both authors leave little room for the possibility that the other's conception might also be correct.²⁴ On the one hand, this is understandable, since the two accounts of virtue are very different and might seem to be incompatible. But again, a more plausible diagnosis suggests itself: namely, that the authors in question are concerned with fundamentally different conceptions of a

²¹ See Axtell (1997).

²² See Zagzebski (1996) and Greco (2002).

²³ The responsibilist conception and a personal worth conception may not be *identical*, but one natural and plausible way of analyzing the notion of personal excellence or worth is in terms of personal responsibility.

²⁴ Driver's main objection to a personal worth conception, for example, is that one can be virtuous without possessing the sorts of motives or intentions that such a conception deems central to virtue. All that matters, she claims, is that the trait in question have good consequences (1996, pp. 116-22). But this of course is entirely consistent with the view that there are *two* kinds of virtue, one of which (viz. a "competence virtue") is a matter of being reliable and does not strictly speaking require the relevant motives or desires, but the other of which does. Slote, on the other hand, analyzes the concept of a moral virtue in terms of certain admirable motives the value of which he says is intrinsic and not derivative from the good consequences to which they are likely to lead (2001, pp. 21-23). However, Slote does not (to my knowledge) acknowledge the possibility that this might be but one of (at least) two viable ways of thinking about the notion of a virtue, the other of which *is* centered on the notion of reliability.

virtue (i.e., Driver with a competence conception and Slote with a personal worth conception), and hence that they are not so much engaged in substantive disagreement as they are talking about different kinds of moral virtue.

The distinction in question also allows us to make sense of the related and (heretofore) seemingly hard-nosed reliabilist claim noted above to the effect that reliability *must* be a defining feature of a virtue, that indeed, it is *the* defining feature. Commitment to this claim, we saw, might lead one to hold that, for instance, Tori and Ted cannot possibly be genuinely virtuous since they utterly fail to achieve their intellectual and moral goals. While false relative to a personal worth conception, this judgment is *entailed* by a competence conception.²⁵ Thus we again may hypothesize that those who find the initial claim convincing and consequently are resistant to the sort of intuitive argument offered above, are thinking in terms of a competence (rather than a personal worth) conception of a virtue.

There are several good reasons, then, for distinguishing between what we have called a “personal worth” conception and a “competence” conception of a virtue. This distinction permits two additional conclusions worth noting. First, it reveals that a *mixed* conception of a virtue makes the mistake, alluded to earlier, of trying to wed two fundamentally different and logically independent conceptions of a virtue under a single, univocal conception. Accordingly, an appeal to a mixed conception cannot provide a way around the principled argument against reliability thesis discussed above. Second, the distinction between a personal worth conception and a mixed conception shows that while reliability is not a defining feature of a virtue qua personal excellence, it is a defining feature – indeed is *the* defining feature – of an alternative conception of a virtue.

²⁵ Or at any rate a competence conception according to which virtues are categorically reliable.

Again it is both necessary and sufficient for the possession of a certain kind of virtue that one possess a competence of the relevant sort.

III

In the remainder of the paper, I want to look in a little more detail at the role of reliability on a personal worth conception of a virtue. One might get the impression from the discussion thus far that reliability is more or less irrelevant to such a conception since it might be thought that all this conception requires for the possession of a virtue is the instantiation of certain admirable aims, motives, desires, etc.²⁶

But this impression is mistaken. We can begin to see why by considering an objection to the effect that the personal worth conception is too lenient. Imagine, for instance, a psychic or fortuneteller who has a genuine desire to reach the truth about the future and who pursues this goal carefully and diligently, but only via her crystal ball. Imagine also an avid sports fan who out of a sincere desire to please his children regularly takes them to various sporting events even though they show little interest or enjoyment in sports. Intuitively, neither of these characters is genuinely virtuous. And yet it might be thought that according to a personal worth conception they would be virtuous since they possess the good motives or desires essential to this conception.²⁷ A solution would be to build the notion of reliability into the personal worth conception (thereby resulting in a mixed conception), since the means these individuals adopt in pursuit of

²⁶ This resembles what Slote refers to as the charge of “autism and antinomianism” against (what is essentially) a personal worth conception of a virtue (2001, pp. 15-19). Slote’s response to the objection, however, is very different from the one developed here.

²⁷ These cases may seem improbable on the grounds that if the psychic *really* had a desire to reach the truth, she would not bother consulting her crystal ball, or that if the father was *genuinely* concerned with his children’s happiness, he would not continue taking them to sporting events. My response is that while the cases are, for the reasons just indicated, unlikely, they nevertheless represent genuine *possibilities*, which is all that is required for the present point.

their virtuous goals seem clearly to be unreliable.²⁸ We have seen, however, that such a requirement is too strong.

Thus the personal worth conception of a virtue is faced with a dilemma. On the one hand, its requirements must be sufficiently strong so as not to deem virtuous the sorts of characters just noted. But on the other hand, its requirements must not be so strong as to make reliability a defining feature of a virtue.

In considering how this challenge might be met, it will be helpful to have at our disposal a distinction between what we might call “virtue-relevant” (hereafter v-relevant) goals and v-relevant actions. V-relevant goals are the goals proper to intellectual and moral virtue. V-relevant actions are those that an agent performs in her attempt to accomplish her v-relevant goals. Thus the fortuneteller’s v-relevant goal is the truth and her v-relevant actions consist in her study of the crystal ball, while the sports fan’s v-relevant goal is the happiness of his children and his v-relevant action that of taking them to various sporting events.

Compare the characters just discussed with Tori and Ted discussed above. The two sets of characters have much in common: both are appropriately motivated by v-relevant goals and yet both perform v-relevant actions that are largely unsuccessful at achieving these goals. Nevertheless Ted and Tori intuitively are virtuous while the fortuneteller and sports fan are not. What then is the difference? It would seem to concern the *cognitive perspectives* of the various characters in question.²⁹ Tori and Ted have

²⁸ This depends of course on how exactly we characterize the means or methods in question. And as the “generality problem” in epistemology shows (see Feldman and Conee 1998), this can be difficult. However, I take it there is a reasonably clear and relevant sense in which the methods in question are *not* reliable.

²⁹ The difference cannot be that Ted and Tori are indexically reliable while the other characters are not. For again, we saw above that intuitively Ted and Tori are virtuous not on account of what they might accomplish in other possible worlds, but on account of the content of their character considered in its own

every *reason to believe* that their v-relevant actions are a perfectly good means to their v-relevant goals, for again, the demons controlling their worlds do seamless work, leaving no indication of any deception or trickery. The fortuneteller and sports fan, by contrast, presumably *lack* good reason to believe that their v-relevant actions are an effective means to their v-relevant goals. Consequently they exhibit a kind of *irrationality*. And this irrationality, it seems, is inconsistent with the kind of personal value constitutive of virtue on a personal worth conception.

This suggests that a personal worth conception can pass through the horns of the dilemma noted above by adopting a certain rationality constraint. It can hold that being virtuous, while in part a matter of possessing certain admirable motives, desires, intentions, etc., also involves the possession of a certain cognitive perspective on one's v-relevant actions. Specifically, it involves having good reason to think that one's v-relevant actions are likely to be an effective means to one's v-relevant goals, that is, that one's v-relevant actions are *reliable*.³⁰

There is, then, a closer connection than one might initially expect between the notion of reliability and a personal worth conception of a virtue. While it is not essential to the possession of a virtue thus conceived that one actually be reliable, it is essential that one have good reason to *think* one is reliable.

right. Moreover, it would be possible to modify the latter cases such that the actions in question clearly are weakly reliable, but where the characters who perform them are less than virtuous. Suppose, for instance, that the father regularly takes his children to parks and toy stores and libraries out of a desire to make them happy. Relative to many possible worlds (our own world included) his v-relevant actions are a reliable means to his v-relevant goal. We can imagine, however, that his kids are atypical and don't enjoy any of these things. If they have given him no indication to the contrary, and indeed have repeatedly told their father that they dislike these activities, while his v-relevant actions would be indexically reliable, he still would not be virtuous.

³⁰ Reliable, of course, in the world in which one finds oneself.

One additional point is worth noting concerning the role of reliability on a personal worth conception. To begin, notice that while I have thus far been arguing that reliability is not an *essential* or *defining* feature of a virtue conceived as a personal excellence (i.e., that it is not part of what *makes* a trait a virtue), I have refrained from claiming that it is not a *necessary* feature. This has not been accidental. Indeed, as I turn now to explain, given certain aspects of some of the key psychological ingredients of a virtue, the traits in question are in fact necessarily indexically reliable. While this fact is not, for reasons already discussed, part of what *makes* these traits personal excellences, it is a necessary concomitant of them nonetheless.

The necessity in question is a function of two main aspects of the various intentions, motives, desires, etc., that (partially) comprise a virtue conceived as a personal excellence. The first of these is their intentional or goal-oriented nature. As we have seen, to possess a virtue is in part to be motivated by, to desire, to aim at, etc., certain valuable ends or goals. The second aspect is what we might refer to as their “robust” nature. It is clear that a person who, say, merely wishes to reach the truth or merely daydreams about a more just society would fail to satisfy the motivational or intentional requirements of a virtue. The same could be said for a person who has only a very weak desire to accomplish these ends. Indeed it is reasonable to think that in order to contribute logically to one’s personal worth or excellence in the relevant way, the qualities in question would need to be sufficiently robust such that, at least with respect to certain environments or situations, they would be *efficacious* in the sense that they would (perhaps together with certain related psychological qualities and some measure of luck) prove an effective means to the goal or end proper to the virtue in question. But this of course is just to say

that when one instantiates the essential intentional ingredients of a virtue conceived as a personal excellence, one necessarily is indexically reliable.

In this section we have reached two additional conclusions regarding reliability and virtue. We have seen that while reliability is not an essential or defining feature of a virtue conceived as a personal excellence, a certain rational *belief* pertaining to reliability is an essential feature. We have also seen that (indexical) reliability itself, though again not a *defining* feature of a virtue thus conceived, nevertheless is a *necessary* feature.³¹

³¹ [Acknowledgements omitted for blind review.]

References

- Aristotle. *Nicomachean Ethics*. Trans. Roger Crisp (Cambridge: Cambridge University Press, 2000).
- Axtell, Guy. 1997. "Recent Work In Virtue Epistemology," *American Philosophical Quarterly* 34: pp. 1-27.
- Axtell, Guy, ed. 2000. *Knowledge, Belief, and Character* (Lanham: Rowman and Littlefield).
- Bentham, Jeremy. 1789. *An Introduction to the Principles of Morals and Legislation*, ed. W. Harrison (Oxford: Oxford University Press, 1948).
- Crisp, Roger, ed. 1996. *How Should One Live?* (Oxford: Clarendon Press).
- Crisp, Roger and Michael Slote, eds. 1997. *Virtue Ethics* (Oxford: Oxford University Press).
- Dancy, Jonathan. 2000. "Supervenience, Virtues, and Consequences," in Axtell 2000: pp. 73-85.
- DePaul, Michael and Linda Zagzebski, eds. 2003. *Intellectual Virtue: Perspectives from Ethics and Epistemology* (Oxford: Oxford University Press).
- Driver, Julia. 1996. "The Virtues and Human Nature," in Crisp 1996: pp. 111-130.
- Feldman, Richard and Earl Conee. 1998. "The Generality Problem for Reliabilism," *Philosophical Studies*, 89: pp. 1-29.
- Garcia, J. L. A. 2003. "Practical Reason and Its Virtues," in DePaul and Zagzebski 2003: pp. 81-107.
- Greco, John. 2002. "Virtues In Epistemology," in Moser 2002: pp. 287-315.
- Hurka, Thomas. 2001. *Virtue, Vice, and Value* (Oxford: Oxford University Press).

- Hursthouse, Rosalind. 1999. *On Virtue Ethics*. (Oxford: Oxford University Press).
- Montmarquet, James. 1993. *Epistemic Virtue and Doxastic Responsibility*. (Lanham: Rowman and Littlefield).
- Montmarquet, James. 2000. "An 'Internalist' Conception of Intellectual Virtue," in Axtell 2000: pp. 135-148.
- Moser, Paul, ed. 2002. *Oxford Handbook of Epistemology* (Oxford: Oxford University Press).
- Nagel, Thomas. 1979. *Mortal Questions* (Cambridge: Cambridge University Press).
- Slote, Michael. 1997. "Agent-Based Virtue Ethics," in Crisp and Slote 1997: pp. 239-262.
- Sosa, Ernest. 1991. *Knowledge In Perspective* (Cambridge: Cambridge University Press).
- Swank, Casey. 2000. "Epistemic Vice," in Axtell 2000: pp. 195-204.
- Williams, Bernard. 1981. *Moral Luck* (Cambridge: Cambridge University Press).
- Zagzebski, Linda. 1996. *Virtues of the Mind* (Cambridge: Cambridge University Press).