13   The situationist critique of virtue ethics

Traditional philosophical theories of virtue define a "virtue" as a species of character trait. Many contemporary philosophical theories of virtue follow suit, though not all do. Adopting this traditional definition exposes a theory of virtue to what has come to be known as the "situationist" critique of virtue ethics. To explain this critique, and to keep track of the ensuing debate, it helps to distinguish philosophical situationism from psychological situationism (see Snow 2010).

Psychological situationists are not philosophers and they make no philosophical claims. Rather, they belong to a particular experimental tradition within social psychology, a tradition that is opposed to traditional personality theory or "personology" (for an accessible introduction, see Ross and Nisbett 1991). Since they are the original situationists, I shall henceforth refer to psychological situationists as "situationists" *tout court*. Philosophical situationists – principally, Gilbert Harman (1999, 2000) and John Doris (1998, 2002) – reject theories of virtue that employ the traditional philosophical definition of 'virtue.' Specifically, they claim that such theories are "empirically inadequate" and their argument for this claim centrally appeals to the experimental results of situationism. It is *their* argument that constitutes the "situationist critique."

The philosophical situationists' argument can be framed as a pair of inferences, which take us from the original situationist data to their own bottom-line conclusion. Their bottom-line conclusion is that traditional theories of virtue are "empirically inadequate."

According to philosophical situationism, the experimental results of situationism demonstrate that most people do not have any character traits. On this basis, philosophical situationists first infer that most people do not have any virtues, as traditionally defined. They then infer, second, that not enough people have any virtues so defined (i.e. not enough for traditional virtue theory to be empirically adequate).

Naturally, this framing of their critique needs some refinement. In particular, we need to specify what kind of character trait is at issue and also to determine how many people are supposed, by traditional virtue theory, to have virtues in the first place (and hence, how many people lacking virtue are "enough" to expose that theory as empirically inadequate). Let us make a start on the first task, postponing the second task for the moment.

At a minimum, we can understand a *character trait* to be a reliable disposition a person has to behave in certain characteristic ways. While plainly requiring further articulation, this initial specification has the merit of satisfying a crucial condition of dialectical adequacy: namely, that specifications of "character trait" be univocal across the philosophical situationists' first inference. For if the specification of "character trait" under which situationism demonstrates that most people do not have any "character traits" differs from the specification under which a virtue is defined as a species of "character trait," the philosophical situationists' critique will fail by equivocation. Later we shall encounter grounds for concluding that their critique does fail in this way.

Situationists actually distinguish two dimensions along which someone's disposition to behave in some characteristic way may be "reliable." One dimension they call "temporal stability," while the other they call "cross-situational consistency." A person has a temporally stable behavioral disposition when she behaves in the same specific way in response to repeated encounters with the same fairly specific situation: for example, when she copies from an answer key on a classroom test in the summer and then copies from an answer key on a classroom test again in the winter. A person has a cross-situationally consistent behavioral disposition when she behaves in the same characteristic way – as distinct, that is, from the same specific way – in response to encounters with a diversity of specific situations, each of which is nevertheless relevant to the

characteristic behavior in question. For example, someone's dis-
position to dishonesty is cross-situationally consistent when she
not only copies from an answer key on a classroom test, but also
pockets some stray change she finds on a classroom desk – assum-
ing, at least, that cheating and stealing are both characteristic of
"dishonesty."[1]

When traditional theorists of virtue define a virtue as a species
of character trait, their definition entails that a virtue includes a
reliable disposition to behave in certain characteristic ways.[2] Thus,
the virtue of courage includes a reliable disposition to behave coura-
geously, the virtue of compassion includes a reliable disposition to
behave compassionately, and so on. Of course, not any old reliable
disposition to behave (say) courageously will express or evidence the
virtue of courage. A reliable behavioral disposition is only a neces-
sary condition of the relevant character trait, not also a sufficient
condition.

What further conditions a person's reliable behavioral disposition
has to satisfy in order to qualify as a virtue depends on the partic-
ular theory of virtue. For example, most theories will not qualify a
reliable disposition to behave "courageously" as a virtue unless the
agent's "courageous" acts are performed *for the right reasons*.[3] Some
theories will not so qualify that disposition unless the agent's "coura-
geous" acts are performed *wholeheartedly*. Other theories will not
so qualify that disposition unless the agent also *has all of the other
virtues*. Aristotle's theory of virtue imposes all three of these further
conditions, but some are more controversial than others.

So traditional theorists of virtue accept that having a reliable
behavioral disposition is a necessary condition of having a virtue.
More specifically, the behavioral disposition they require has to be
reliable along *both* of the dimensions situationists distinguish. In
other words, a person's reliable disposition to behave (say) compas-
sionately has to be temporally stable *and* cross-situationally con-
sistent. As we shall see, the focus of the debate between philo-
sophical situationists and traditional theorists of virtue concerns
the cross-situational consistency of the relevant behavioral dispo-
sitions. Philosophical situationism's fundamental empirical claim
is that most people do not have any cross-situationally consistent
behavioral dispositions. If most people lack any cross-situationally
consistent disposition to behave compassionately, it certainly

follows that most people lack the virtue of compassion, as traditional theorists of virtue define it.

We should notice a final preliminary point about the meaning of "character trait." In principle, a behavioral "disposition" is importantly distinct from a mere behavioral "regularity." Among other things, the former goes beyond the latter in offering a particular explanation for it. If someone copies from an answer key on a classroom test every week, his behavior exhibits a certain regularity – specifically, a temporally stable regularity. Strictly speaking, however, this regularity is not sufficient to establish that the person has a temporally stable behavioral disposition. For, in itself, a temporally stable behavioral regularity is consistent with various explanations, whereas a temporally stable *disposition* entails that the corresponding regularity is explained, in particular, by features of the person's individual psychology (rather, say, than by factors *external* to his psychology – e.g. the company he keeps). Since a "character trait" is specified in terms of a reliable behavioral disposition, understood in this more robust sense, attributions of character traits inherit that additional explanatory burden.

In practice, though, the distinction between behavioral regularities and behavioral dispositions has not played a prominent role in the debate because what situationists really deny is the more basic proposition that most people's behavior exhibits any cross-situational regularity. If that were true, there would be nothing for any cross-situationally consistent disposition to explain.

PHILOSOPHICAL SITUATIONISM'S BEST DATA

Let us now consider the evidence philosophical situationists adduce for their fundamental empirical claim that most people lack any cross-situationally consistent character trait. The best place to begin is with Hartshorne and May's (1928) famous study of honesty in children. Hartshorne and May observed how thousands of school children behaved in various specific "honesty-relevant" situations. For example, their observations included a "stealing" situation (some change has been left on a table in an empty classroom and there is an opportunity to take it); a "lying" situation (another child is going to get in trouble and there is an opportunity to avert this by making a false report); and a "cheating" situation (one is correcting

one's own test sheet in class and there is an opportunity to amend one's answers first with the benefit of the answer key). These situations are relevant to the assessment of someone's honesty because we ordinarily expect that an honest person can be relied upon not to steal, not to lie, and not to cheat. In particular, then, we might expect an honest subject in Hartshorne and May's experiment not to pocket the change *and* not to make a false report *and* not to amend her answers from the key.

What Hartshorne and May found, however, was that the average correlation between their subjects' not pocketing the change and their not making a false report was only 0.13; the average correlation between not pocketing the change and not copying from the answer key was also only 0.13; and the average correlation between not making a false report and not copying from the answer key was only 0.31. Overall, the average correlation between any two of Hartshorne and May's behavioral measures of honesty was only 0.23. By contrast, the average correlation between not copying from the answer key on one occasion and not copying from another answer key six months later was 0.79.

In other words, not much cross-situationally consistent honest behavior was observed among Hartshorne and May's subjects, at least not as far as the performances of their "average subject" were concerned. On the other hand, *within* a given "honesty-relevant situation" (such as the cheating situation), there was a good deal of temporally stable honest behavior. To put their findings yet another way, *few* of Hartshorne and May's subjects turned out to be cross-situationally consistently honest, even though many of them exhibited temporally stable behavior in some or other sub-subdepartment of honesty (not even "not cheating," e.g. but only "not copying from an answer key").

Hartshorne and May's data are the best place for us to begin a review of the situationist evidence, for three main reasons. First, their study investigates a character trait that plainly counts as a standard virtue, unlike many other traits of interest to social psychologists, such as talkativeness or dependency.

Second, and most important, the quantitative values they report are perfectly representative of the findings of other explicit investigations of cross-situational consistency. By an "explicit investigation of cross-situational consistency," I mean an experiment in which

subjects are placed in a *plurality* of *different* specific situations, each of which is designed to elicit a manifestation of the same character trait. Summarizing this experimental literature, Ross and Nisbett (1991, p. 95) declare that

the average correlation between different behavioral measures specifically designed to tap the same personality trait (for example, impulsivity, honesty, dependency, or the like) was typically in the range between .10 and .20, and often was even lower... Virtually no coefficients, either between individual pairs of behavioral measures or between personality scale scores and individual behavioral measures, exceeded the .30 "barrier."

This correlational "barrier" represents the heart of the situationist's empirical case.

Third, as a bonus, Hartshorne and May also placed their subjects in iterations of the *same* specific situation, which is what allowed them to collect explicit data on temporal stability. While the correlation of 0.79 for their cheating situation seems to be at the high end of the typical range for stability correlations – Ross and Nisbett describe this range as "often exceed[ing] .40, sometimes reaching much higher" (1991, p. 101) – what remains representative about Hartshorne and May's data is the markedly unfavorable contrast between consistency correlations and stability correlations. This contrast is the reason why situationists focus their skepticism on the cross-situational consistency of behavior, rather than on the reliability of behavioral dispositions quite generally.

WHAT IS THE SCOPE OF A THEORY OF VIRTUE?

So the heart of situationism's empirical case consists in the low ceiling on correlations reported from explicit investigations of cross-situational consistency. Later we shall have occasion to consider the rest of the case. But let us first examine what follows from its heart, focusing again on Hartshorne and May's honesty study.

Hartshorne and May found that most of their subjects were *not* cross-situationally consistent across their behavioral measures of honesty. As we have seen, two inferences lie between this observation and the philosophical situationist's bottom-line conclusion. Each of them turns out to be objectionable. But let us begin with the

second inference. Suppose that most of Hartshorne and May's subjects (therefore) lacked the virtue of honesty and that their findings are widely replicable.[4] Does it follow that "not enough people have the virtue of honesty," as the philosophical situationists' critique requires?

To examine this question, we have to return to the task we left trailing earlier, concerning the number of people who are supposed, by traditional virtue theory, to have any given virtue, such as honesty. In effect, this is a question about the scope of the theory. To whom does a theory of virtue apply (and *how* does it apply to them)? To license her inference from "most people lack the virtue of honesty" to "not enough people have that virtue," the philosophical situationist has to interpret traditional virtue theory as applying to *most* people *and* as applying to them by claiming, among other things, that *they have* the virtue of honesty. However, as many defenders of traditional virtue theory have pointed out, this interpretation is tendentious. To illustrate, let me discuss two alternative interpretations of a traditional theory's scope. One is more subtle than the other, but on neither does it follow that "not enough people have the virtue of honesty." Moreover, these alternatives are consistent, so the virtue theorist does not have to choose between them.

On the first interpretation, virtue theory applies to "most" people (indeed, to everyone), but no claim is made that most people *actually have* the virtue of honesty. Instead, the traditional virtue theorist's claim will be that everyone *should* have that virtue. Her claim will be normative, rather than descriptive. Of course, even this normative claim has at least one empirical presupposition – namely, that for any given person it is possible that he or she can have the virtue of honesty. On the face of it, however, this presupposition is fully consistent with the "fact" that most people lack the virtue of honesty. Indeed, it is arguably consistent with a state of affairs in which no one actually has the virtue of honesty.

Doris and Stich (2005, §2) concede that the empirical evidence does not show that the acquisition of virtue is impossible. They reply that it is possible that the acquisition of virtue is psychologically impossible and that the onus is now on the virtue theorist to demonstrate that it is psychologically possible. If "psychologically possible" means "possible for a real human being [as distinct from merely logically possible]," this is a curious reply. For, as we shall

elaborate next, Hartshorne and May's data are consistent with *some* people's actually having the virtue of honesty (see Adams 2006-[1], p. 116); and from the fact that some people actually have the virtue, it follows that it is possible for a real human being to have the virtue.

But perhaps 'psychologically possible' is meant to gesture at some more demanding empirical constraint on an ethical ideal's having normative purchase on the average human being – more demanding, that is, than mere "possibility for a real human being." Perhaps, for example, it is meant to suggest that an ideal of virtuous character has normative purchase *only if* the average human being has a fighting chance of actually acquiring the relevant character trait (see Doris 1998, p. 525 n. 41). To vindicate this suggestion, and thereby turn the gesture into a reply, would require two arguments, each anchored in the same construal of the average person's "fighting chance." One argument is needed to show that the low ceiling on cross-situational consistency correlations deprives the average person of her fighting chance and a second argument is needed to show that it is philosophically tenable to subject ethical ideals to the resultant empirical constraint (i.e. to water them down that far). The trouble is that the easier one makes it to complete one of these arguments, the harder it will be to complete the other.

For example, suppose we say that the average person has a fighting chance to acquire a virtue only if it is easy for her to acquire it. It will then be simple to read the situationist data as showing that the resultant empirical constraint is not satisfied. However, it will be very difficult to vindicate the corresponding philosophical proposition that "an ideal of virtuous character has normative purchase only if it is *easy for the average human being* to acquire the relevant character trait."

Let us now consider a second alternative. On this interpretation, the only claim that a theory of virtue makes about anyone's actually having the (full measure of) the "virtuous cross-situationally consistent character trait of honesty" is a claim about "some" people, namely, the few exemplars or models of honesty. Accordingly, as long as a small number of people have the virtue of honesty, then enough people have it to satisfy the empirical commitments of the theory. Notice that although the scope of a theory of virtue will be (in one way) narrower here, as compared to the previous alternative, its

empirical exposure will be greater, since the theory will now claim that some people actually have the virtue of honesty.

Still, this claim is entirely consistent with Hartshorne and May's data.[5] While the average correlation of 0.23 between any pair of their behavioral measures of honesty plainly excludes there being many subjects whose own *individual* average correlation (over any pair of the honesty measures) was much higher than 0.23, it is consistent with there being some subjects whose individual average correlation was much higher than 0.23. It is consistent, in other words, with there being a small number of individuals who were cross-situationally consistently honest. Hence, from Hartshorne and May's overall average cross-situational consistency correlation of 0.23, it does not follow that "no one has the virtue of honesty" or even that "next to no one does." But then, on this second interpretation, neither does it follow that "not enough people have the virtue of honesty."

It would be natural to object – as, in effect, Doris (1998, pp. 511–13) does object – that a theory of virtue whose scope is restricted to a few exemplars is not of much interest, either practically or theoretically. To command our attention, a theory of virtue must be shown to have some clear relevance to most people (better yet, to everyone). This much is certainly difficult to deny.

Fortunately, there is also no need to deny it, as help on this point is available in a surprising corner. Recall that some contemporary theories eschew the traditional definition of virtue as a species of character trait. They define virtue instead as a kind of occurrent state – in the simplest case, as a kind of occurrent act. Hurka (2006) argues that occurrent act theories of virtue are immune to the situationist critique. To see why they are immune, consider an occurrent requirement to act honestly – for example, the requirement not to cheat on this occasion.[6] The normative purchase of this requirement does not depend on any possibility of the agent's having or acquiring any disposition to perform acts of the relevant kind (however this kind is construed). Empirically, it depends only on the possibility of the agent's performing the occurrent act, i.e. of his not cheating then and there. Since none of its data casts any doubt on this possibility, occurrent act theories of virtue are unscathed by situationism.

It is common to treat occurrent act theories of virtue and traditional theories as diametrically opposed to each other; and for some

purposes, they are.[7] However, there is also a core component of occurrent act theories to which traditional theories can, and should, help themselves. I have in mind the conception of (individual) virtues as sources of occurrent act requirements on individuals. On this conception, the very existence of a normative requirement not to cheat on this occasion (say) constitutes part of the normative relevance of the virtue of honesty, since "honesty" is itself the source of the requirement (and of kindred others).[8] This highlights a significant aspect of the normative relevance of the virtues, an aspect that not only applies to everyone, but is also immune to the philosophical situationist's critique.

A traditional theory of virtue should incorporate this conception of individual virtues. It should do so, first and foremost, because that conception is anyhow correct. But incorporating that conception would also spell out one obvious way in which the virtues remain normatively relevant to everyone, even if only a few exemplary individuals are expected fully to acquire the corresponding virtuous character traits.

Now it may be further objected that this reply does nothing to rescue the traditional definition of virtue, since it merely conjoins to the traditional theory some free-floating truths (albeit, about virtue) that have nothing to do with its hallmark definition. But that would be a mistake, for the truths in question can readily be made to have something to do with the traditional definition of virtue as a species of character trait. The truths about the specific occurrent act requirements issued by a given virtue will have a firm footing in that hallmark definition, rather than floating free from it, as long as a traditional theory reserves some privileged role for the exemplars of that virtue in identifying those same requirements.[9] A well-known, albeit extreme example of such a privileged role is given by Hursthouse's (1999) claim that *all* right action (and, a fortiori, the occurrent requirements issued by any particular virtue) is *defined* in terms of how a virtuous agent would characteristically act in the situation. But less extreme privileged roles for exemplars are also possible, not to mention much more plausible.

Further to articulate the second alternative, then, a theory of virtue so interpreted will make different claims with different scopes. On the one hand, it will (only) claim that "some" people actually have the virtuous cross-situationally consistent character

trait of honesty. On the other hand, it will also claim that "every-
one" is required to act honestly, occasion by occasion, whether he
or she has the character trait or not. Both claims are consistent with
situationism. They are bridged by the distinctive additional claim
that the exemplars of honesty have a privileged role in identify-
ing the nature of the acts occurrently required by honesty (i.e. in
identifying what it takes to act honestly on some occasion or
another), thereby relating everyone *indirectly* to the virtuous char-
acter trait of honesty.[10] It is this distinctive bridging claim that qual-
ifies the resultant theory as a "traditional" theory of virtue.

WHICH "CHARACTER TRAITS" ARE THE RIGHT ONES?

Let us now set the issue of scope aside and turn to philosophical
situationism's other objectionable inference. We can isolate it by
taking the fact that most of Hartshorne and May's subjects were not
cross-situationally consistent across Hartshorne and May's behav-
ioral measures of honesty, and asking whether it really follows that
most of their subjects lacked the virtue of honesty. To draw that
conclusion, one has to assume that Hartshorne and May's behav-
ioral measures properly operationalize the virtuous character trait
of honesty. Yet that assumption is objectionable.

To operationalize a cross-situationally consistent character trait,
experimenters have to specify a variety of concrete situation–
response pairs. In the case of honesty, the situations have to be "hon-
esty eliciting," that is, situations likely to provoke a response that
can be readily evaluated as "honest" or "not honest." Hartshorne and
May's honesty-eliciting situations include the stealing, lying, and
cheating situations described earlier. For each honesty-eliciting sit-
uation, the experimenters also have to specify the particular response
that counts as the "honest" response. In the cheating situation, for
example, Hartshorne and May specify "not copying" as the honest
response.

We can distinguish three separate respects in which Hartshorne
and May's concrete situation–response pairs fail to operationalize
honesty properly. Philosophically, the most important of these is
that their behavioral measures do not take account of the nor-
mative sensitivity of a virtuous character trait's responsiveness to
situations. Unlike with some kinds of character trait, the responses

characteristic of a virtuous trait do not simply respond to the situation as such. Rather, they respond to some reason for action present in the situation; alternatively, they respond to the normative requirements imposed by some value at stake in the situation (e.g. the value of honesty).

Reasons for action, however, can be neutralized – defeated, cancelled, preempted – by (other) features of a concrete situation, even if those features do not change the identity of that situation under some abstract or mechanical description. Hence, two concrete situation–response pairs may appear equivalent – from a certain point of view, anyhow – despite the fact that the relevant reason for action is operative in only one of them. But the adequacy of a given concrete pairing as a behavioral measure of honesty depends on the reason to act honestly remaining operative there. Otherwise, the concrete situation does not call for any particular response by the agent after all, at least not as far as honesty is concerned.

Hartshorne and May's lying situation illustrates the difficulty well. Ordinarily, making a false report counts both as "lying" and as "trait-contrary" behavior for honesty. A situation presenting an opportunity to make a false report is therefore plausibly regarded as an honesty-eliciting situation and "reports falsely" is plausibly counted as the "not honest" response. Still, like any ordinary moral reason, the reason not to make a false report can be defeated; and arguably, in Hartshorne and May's lying situation, it *is* defeated. For recall that, in their scenario, the false report serves to prevent another child from getting into trouble, i.e. to accomplish some (sufficient) good. In that case, reporting falsely is *not* trait-contrary behavior for honesty – it does not contraindicate possession of the virtuous trait – because the *reason* for action to which the virtuous trait responds is not operative in the situation. So the lying situation is not an adequate behavioral measure of honesty.

A second respect in which Hartshorne and May's behavioral measures are inadequate is perfectly generic, in the sense that it does not stem from anything particular to the virtues as a subset of character traits. In fact, the objection is best explained by reference to one of the fundamental tenets of social psychology, which situationists themselves emphasize greatly. Ross and Nisbett call it the "principle of construal":

The impact of any "objective" stimulus situation depends upon the personal and subjective meaning that the actor attaches to that situation. To predict the behavior of a given person successfully, we must be able to appreciate the actor's construal of the situation. (1991, p. 11)

Since the attribution of a character trait is meant to enable (or at least, to facilitate) predictions of how the bearer of the trait will behave, it seems that the concrete situation–response pairs that operationalize the trait ought to be pairings whose significance is agreed on between the experimenter and the subject (i.e. between predictor and predictee). For example, it seems that the evaluation of specific responses as either "honest" or "not honest" ought to be so agreed on. However, Hartshorne and May used "objective" behavioral measures, meaning that the specification of the particular situations and responses were fixed by the experimenters alone.

To illustrate the resultant difficulty, consider their stealing situation. Some change has been left on a table in an empty classroom and there is an opportunity to take it. Hartshorne and May count taking the change as "stealing," i.e. as the "not honest" response. But someone who believed in "finders keepers" – call her Sally – would disagree. Sally would find nothing wrong with taking the change and, more significantly, nothing inconsistent in both "not copying from the answer key" and "taking the change." Say that is indeed how she responds in the cheating and stealing situations. In scoring Sally as "cross-situationally *in*consistent," Hartshorne and May are really (only) registering their disagreement with her over the correctness of finders keepers, rather than discovering any true behavioral inconsistency on Sally's part. Furthermore, since they disregard her belief in finders keepers, it should come as no surprise that Hartshorne and May would have had trouble predicting Sally's behavior in the stealing situation.

Of course, *morally*, it may well be that Hartshorne and May are correct and that finders keepers is not a valid principle. Let us stipulate that it is invalid. In that case, Sally will turn out to be behaviorally consistent, but also not (fully) honest. Recall that behavioral reliability – and, a fortiori, cross-situational consistency – is only a necessary condition of virtue and not also a sufficient condition. To the other candidate necessary conditions mentioned earlier, we can now add a further one: possession of the correct moral beliefs

(e.g. beliefs about the validity of finders keepers).[11] Since Sally has false moral beliefs about honesty, she cannot be a fully honest person, no matter how consistently she behaves. Likewise, whether Hartshorne and May sincerely agree with Sally about finders keepers or not has no bearing whatever on the validity of the principle itself. The two necessary conditions (along with their subcomponents) operate independently.

Hartshorne and May's behavioral measures of honesty therefore fail to satisfy a condition of adequacy on behavioral measures of any character trait, namely, that experimenter and subject agree on the significance of the relevant concrete situation–response pairs.[12] This condition is imposed by the need to license predictions of the subject's behavior and has no consequences for the correctness of any moral belief about the particulars of what honesty calls for.

A final respect in which Hartshorne and May's behavioral measures of honesty are inadequate arises from the fact that relevance – specifically, the relevance of a given honesty-eliciting situation to the normative requirements imposed by the value of honesty – is a matter of degree. We can plausibly distinguish between marginal and paradigmatic cases – or, to borrow from psychological idiom, between highly and not highly prototypical cases – of either "honest" behavior or "not honest" behavior. In this vein, copying from an answer key does seem like a paradigmatic case of cheating. By contrast, pocketing stray change is only a marginal case of stealing, even if one rejects the "finders keepers" principle. Much better examples of stealing would be shoplifting (children) or purse snatching (adults). Morally and psychologically, the expectation of cross-situational consistency is much stronger across situations that elicit paradigmatic cases of honest behavior. Insofar as adequate behavioral measures of honesty should accordingly be confined to paradigmatic cases, Hartshorne and May's stealing situation is not an adequate measure.

ONE-TIME PERFORMANCE EXPERIMENTS

A low ceiling on cross-situational consistency correlations, then, does not entail that traditional theories of virtue are "empirically inadequate." As we have seen, there are at least two gaps on the path from "most people lack character traits" to "not enough people

have any virtuous character traits." To begin with, "most people" in the premise does not include enough people to warrant the "not enough people" in the conclusion. In addition, the "character traits" operationalized in the premise are not the right kind to contradict the "virtuous character traits" in the conclusion. Sometimes (normative sensitivity) that is because they do not specifically qualify as "virtuous." Other times (construal, paradigmatic relevance) it is simply because their operationalization is generically defective.

However, as even casual acquaintance with the writings of philosophical situationists makes clear, what I have called the heart of their empirical case – the 0.30 correlational ceiling – is only one thread among many in the tangle of evidence they adduce.[13] At least as much attention is devoted to studies of a rather different sort from Hartshorne and May's, of which Milgram's obedience to authority experiments, Darley and Batson's Good Samaritan experiment, Isen and Levin's dime in the phone booth experiment, and Latané and Darley's bystander experiments are prominent examples. In these other experiments – of which there are "hundreds, if not thousands," as Doris is often at pains to emphasize – a majority of subjects (sometimes more) are not only led to perform trait-contrary behavior, but are apparently so led by a trivial feature of the situation in which the subjects have been placed. What is to be made of this part of the philosophical situationist's case?

As the first order of business, we should distinguish various conclusions that might be drawn from this other evidence. For present purposes, I shall limit myself to two possibilities. On the one hand, there is the proposition that *most subjects lack some cross-situationally consistent character trait*, which I shall call the "official interim conclusion." On the other hand, there is the proposition that *situational variables have a powerful effect on behavior*, which I shall call the "situationist minimum." Let us start with the former.

The official interim conclusion is only an *interim* conclusion because, as we know, even in the best-case scenario, some distance remains between it and the philosophical situationist's bottom-line conclusion, "not enough people have any virtues." But let us now ignore that distance and the previously observed impediments to traversing it. What is more useful to notice here is that the official interim conclusion does clearly follow from Hartshorne and

May's study. In their case, the cross-situationally consistent character trait that most of their subjects lacked was "honesty$_{HM}$," where the subscript indicates Hartshorne and May's particular operationalization of honesty. Functionally, in other words, the official interim conclusion captures precisely the contribution to the philosophical situationist's bottom line made by Hartshorne and May's study. In examining whether the rest of the philosophical situationist's case also licenses her official interim conclusion, we are therefore considering whether this other evidence is on a par with Hartshorne and May's.

To make efficient progress on this question, let me introduce a fundamental structural distinction between experiments like Hartshorne and May's and experiments like Milgram's or Darley and Batson's. The distinction holds between experiments structured as an "iterated trial" and those structured as a "one-time performance." In an iterated trial experiment, each subject is placed in a plurality of eliciting situations, whereas in a one-time performance experiment each subject is placed in a single eliciting situation only. We can actually further distinguish two different species of iterated trial design, depending on whether the "plurality" of situations is constituted by simple iterations of the original situation or by a diversity of specific situations designed to elicit manifestations of the same trait. The second species of iterated trial design thus corresponds to what I earlier called an "explicit investigation" of cross-situational consistency.

Hartshorne and May's study is an iterated trial experiment (indeed, it instantiates both species of this design at once). What characterizes the other experiments we are examining is that they are all one-time performance experiments.[14] In Milgram's famous experiment, for example, each subject is presented with a single incremental sequence of opportunities to "punish" a learner under the instruction of an authority figure (the experimenter); and the aim is to discover at what point (if any) the subject will disobey.

The question before us, then, is whether any one-time performance experiment can license the (philosophical situationist's official interim) conclusion that most subjects lack some cross-situationally consistent character trait. Consider Milgram's experiment. No matter where in the sequence one plausibly draws the

line past which the virtue of compassion required his subjects *not* to proceed, most of them crossed the line. So from the standpoint of compassion, the behavior of most of Milgram's subjects was plainly trait contrary. Does it follow that they lacked the cross-situationally consistent trait of compassion?[15]

It all depends on how high we set the degree of reliability required of a trait.[16] If we define a cross-situationally consistent trait as an exceptionless behavioral disposition (i.e. as 100-percent reliable), then it certainly follows that Milgram's subjects lacked the trait of compassion. However, when applied to virtue, that would yield an implausibly strict definition. It would identify virtue with nothing less than utter perfection, "one strike and you are out." In any case, Doris explicitly declines to define his notion of a "trait" as an exceptionless disposition (2002, p. 19). So we can set that definition aside.

Unfortunately, on any other setting, it no longer follows that Milgram's subjects lacked the trait of compassion. In other words, when any lower (and hence, more plausible) degree of reliability is used to define the trait, our license to draw the official interim conclusion disappears. The basic difficulty is that an ordinarily reliable behavioral disposition, even if it is very reliable, is thoroughly consistent with a single episode of "contrary" behavior. To exclude possession of a(n imperfectly) reliable behavioral disposition, we therefore require *repeated* failures.[17] Yet since Milgram's obedience experiment is a one-time performance experiment, his subjects never get the chance to fail repeatedly. By definition, all one-time performance experiments share this fatal defect.

Now it may be tempting, in reply, to suppose that a subject's one-time violation of the requirements of compassion is still best explained by her *lacking* the cross-situationally consistent trait of compassion, even though her violation is logically consistent with her possession of that trait.[18] But however tempting, and however convenient for the philosophical situationist, this would be a mistake. In fact, it would be very close to the same mistake situationists themselves diagnose under the label of the "fundamental attribution error" (Ross and Nisbett 1991, chap. 5). The common element in both mistakes is the tendency to decide the possession of a behavioral trait (either for or against) on an insufficient evidence base – typically, on the basis of a single observation, which is all that a

one-time performance experiment can ever yield. Both mistakes are forms of jumping the evidential gun.

By contrast, a true warrant for the attribution or exclusion of a reliable behavioral trait requires a *plurality* of observations of the subject's behavior in trait-relevant eliciting conditions. In the particular case of a cross-situationally consistent trait, as distinct from a temporally stable trait, the plurality of observations has furthermore to be *distributed* over a variety of trait-relevant eliciting conditions. We thereby return to the point that genuine evidence for the (non-)existence of cross-situationally consistent character traits is properly to be sought from the kind of experiment exemplified by Hartshorne and May's study – that is, from at least the second species of iterated trial experiment.[19]

SALVAGE OPTIONS

So one-time performance experiments cannot license the official interim conclusion that most subjects lack some cross-situationally consistent character trait. For present purposes, that leaves us with the situationist minimum as an alternative, i.e. with the conclusion that situational variables have a powerful effect on behavior. Let us simply accept that this conclusion is well supported by the great number of one-time performance experiments on record, as the pertinent questions here do not concern this basic inference. They are, rather, "Just *how* powerful is this effect?" and, more so, "By what argumentative route(s) can the situationist minimum be connected to the philosophical situationist's bottom-line conclusion that not enough people have any virtues?"

I do not know the answer to the first question. But I shall sketch two specific routes one might pursue in answering the second. Since neither will prove compelling, it is also worth saying that it is really the philosophical situationists who owe us an answer to the second question. On the whole, they have been remarkably casual on this point (see Doris 2002, p. 38), especially considering that it is not at all obvious how exactly one is supposed to get from the situationist minimum to their bottom line.

To preserve a measure of tidiness along the first route, let us restrict our attention to one-time performance experiments in which the experimental situation is compassion eliciting. Now suppose we

assume that, for any given one-time performance experiment, most people outside the actual subject population *would have* behaved as the *average* subject did in fact behave, had these people somehow found themselves in the relevant experimental situation (either in an experiment or in real life). In addition, suppose we further assume that our first assumption *continues to hold*, even as we go on counterfactually "inserting" a given population into *each and every* one-time performance experiment in our restricted universe. Finally, let us take ourselves to be entitled, on this basis, to conclude that most people have a significant number of compassion-contrary performances on their "composite moral record," that is, on the moral record composed of their actual and counterfactual behavior.

Let me call this the *counterfactual addition* route. Since it is manifestly a loose construction, I shall not pause to examine its various component steps. What I wish to emphasize instead is how far we nevertheless remain from the philosophical situationist's bottom line. If most people have a significant number of compassion-contrary performances on their composite moral record, then we may well be able to subtract a significant margin from the maximum degree of "overall reliability" that most people are able to achieve in relation to compassion. (This notion of "overall reliability" is a generic amalgam of temporal stability and cross-situational consistency, and hence distinct from either.[20] We are forced to resort to it because the counterfactual addition route tells us nothing about the distribution of a person's trait-contrary performances over suitably individuated kinds of eliciting situation.[21])

Nothing follows from this, however, about most people's maximum degree of cross-situational consistency in relation to compassion. Nothing follows precisely because overall reliability is distinct from cross-situational consistency. As a result, the counterfactual additional route fails to deliver the philosophical situationist's official interim conclusion (which is about cross-situational consistency).[22] But while this may be worth noticing, it need not be fatal. For one thing, it might be argued that some minimum degree of "overall reliability" in behavior is also a necessary condition of virtue.

By contrast, a further limitation is crippling: the counterfactual addition route is subject to the same scope objection that afflicts the official interim conclusion itself. The fact that "most people's" maximum overall reliability in relation to compassion is capped in some

significant way tells us nothing about the overall reliability that can
still be achieved by a "few" exemplars of virtue: neither what the
most reliable compassionate person's maximum is nor how many
people have achieved that maximum of overall reliability in com-
passion. Hence, the counterfactual addition route cannot yield the
conclusion that not enough people have the virtue of compassion.

A second route to the philosophical situationists' bottom line
attempts both to deflect the scope objection and to bypass their offi-
cial interim conclusion. It begins by taking seriously the bewildering
number and open-ended variety of adverse situational influences on
moral behavior (see Sabini and Silver 2005, p. 545). In any given situ-
ation calling for a virtuous performance, everyone is exposed to this
corrosive array of adverse influences. Moreover, *anyone* (including
an exemplar of virtue) can succumb to one of them, and thereby
fail to do the virtuous thing. No one with this kind of vulnerabil-
ity, the argument continues, can really be "relied" upon to behave
virtuously, not in the sense that the traditional conception of virtue
demands anyhow. Since everyone has that vulnerability, no one has
any virtues. Let me call this the *extreme moral randomness* route.

Suppose we swallow this argument (hook, line, and sinker). It
might well be resisted, but let us not. For all that, an important dif-
ficulty remains. To wit, the extreme moral randomness route proves
too much and also proves too little. It proves too much insofar as
it fails to distinguish virtue ethics from its relevant rivals. In terms
of rivals in moral theory, the threat in question – namely, that of
a moral agent's being overwhelmed on a given occasion by a situa-
tional influence – applies, not only to an agent trying to be virtuous,
but to any agent trying to follow *any* ordinary moral rule (see Sabini
and Silver 2005, p. 536 n. 5). Accordingly, if this threat is a problem
for virtue ethics, then it is equally a problem for any nondebunking
conception of morality (and, a fortiori, for deontology, consequen-
tialism, or what have you).

In terms of rivals in trait psychology, the same threat operates
just as well against an agent with a temporally stable trait as it does
against an agent with a cross-situationally consistent trait. After all,
the adverse influences are precisely situational. They operate within
the confines of a given situation and their effectiveness is presum-
ably independent of whether the agent's behavior on that occasion
is being compared with her behavior on narrowly similar occasions
or with her behavior on trait-relevantly similar occasions.[23] Thus,

while Doris (2002, pp. 23, 64) affirms the existence of temporally stable traits, the extreme moral randomness route is actually inconsistent with them.

Granted a modest clarification, the extreme moral randomness route also proves too little. As initially presented, the argument is (deliberately) vague about just how "powerful" it claims the array of adverse situational influences to be.[24] The modest clarification specifies that these situational influences are *not so* powerful as to undermine the agent's ordinary moral responsibility for his behavior, which entails that it is still *possible* for him to do the right or virtuous thing on the occasion. As we saw above, however, the possibility of an agent's performing the act that virtue requires of him is enough to secure the normative purchase of the occurrent requirement he faces in the situation. In turn, the fact that the occurrent act requirements issued by a given virtue have normative purchase is enough to establish the normative relevance of that particular virtue. Although none of this was news for an occurrent act theory of virtue (see Hurka 2006), we learned that a traditional theory of virtue can help itself to this path to normative relevance, too. Yet, insofar as the extreme moral randomness route is consistent with the normative relevance of a traditionally defined virtue, it proves too little.

Of course, the philosophical situationist can always reject this clarification, modest or not. But that would make it all the more clear that the extreme moral randomness route proves too much, since its argument would then lead to a form of moral nihilism. Indeed, we can now refashion our difficulty as a dilemma. Either the extreme moral randomness route accepts the modest clarification or it does not. If it does, it is consistent with the normative relevance of the virtues, even as traditionally defined, and so proves too little. If it does not, it leads to moral nihilism, which undermines all nondebunking moral theories, and so proves too much.

CONCLUSION

Since the situationist critique of virtue ethics is an empirical critique, its fate – like that of any empirical question or set of questions – can ultimately only be decided by empirical evidence. This means that there is only so much we can conclude about it on

the basis of analysis alone. It also means it may yet turn out that virtue ethics is empirically inadequate. None of the objections to the situationist critique we have canvassed can inoculate virtue ethics against that possibility (nor were they intended to).

As things stand, we either lack the evidence or we lack the argument(s) to vindicate the philosophical situationist's bottom-line conclusion that not enough people have any virtues. On the evidence side, what would be most directly relevant would be evidence from iterated trial experiments (see note 19) that told us roughly how many people (in the population at large) have behavioral traits belonging to a privileged subset of cross-situationally consistent character traits. Leaving aside the issue of which traits map on to which specific virtues, the precise boundaries of the privileged subset depend on the structure of one's favored traditional theory of virtue (as well as on various moral truths). But even to get in the right general neighborhood, the cross-situationally consistent traits have to be normatively sensitive and their operationalization has to incorporate certain basic points of agreement between the subject and the experimenter. These are necessary conditions of membership in the privileged subset, whatever the details of one's favored theory. Moreover, they are conditions not satisfied by Hartshorne and May's landmark honesty study, which is the best existing evidence available to the philosophical situationist.

On the argument's side, there may be other ways to reach the conclusion that "not enough people have any virtues" from the point that situational variables have a powerful effect on behavior, beyond the few considered here. Despite what I have suggested, then, it may be possible to rest the philosophical situationist's critique on the basis of evidence from one-time performance experiments after all. Finally, of course, it is also possible that some of the arguments I have advanced in defense of virtue are simply mistaken – perhaps, for example, there is some compelling objection to (both versions of) the scope objection, the force of which philosophical situationists can help us to appreciate.

NOTES

1 The example also assumes that copying from an answer key is "cheating" and that pocketing stray change is "stealing."

2 For simplicity, let us understand "traditional theorists of virtue" to
   include contemporary theorists who adopt the traditional definition of
   virtue. This will free us from always having to acknowledge that some
   contemporary theorists of virtue (e.g. Hurka 2001) do not define a virtue
   as a species of character trait. We shall return to the significance of that
   theoretical option below.

3 But some explicitly omit this condition (e.g. Driver 2001). Driver also
   rejects the other two conditions to follow in the text.

4 One reason to question this, of course, is that Hartshorne and May's
   subjects were all children (see e.g. Kamtekar 2004, p. 466 n.). But I shall
   put this point to the side, in order to concentrate on the more general
   issues at stake.

5 The most we can say here concerns what is *consistent with* their data,
   since their data do not track the behavior of individuals across situa-
   tions. Rather, they are reported at the level of a population aggregate.
   Statements about individuals – e.g. that some of them had high average
   cross-situational consistency correlations – therefore have to be inferred
   from aggregate comparisons.

6 For convenience, I shall speak in terms of requirements to act virtu-
   ously. But this can be read compatibly with various views about the
   moral weight or precise deontological status of what virtue calls for in
   a given situation.

7 Centrally, there is an important question about the priority of acts
   versus dispositions about which the two kinds of theory essentially
   disagree. For some discussion, see Thomson (1997).

8 When referring to honesty as the source of occurrent act requirements, it
   would be better to distinguish the "value" of honesty from the "virtue"
   of honesty, where the value informs the corresponding virtue. This
   makes it clear that no metaphysical priority need be claimed for the
   virtuous exemplar, who can remain metaphysically independent of the
   source of the requirements to which she responds.

9 Following the distinction registered in note 8, this privileged role can
   be understood as epistemically privileged (and, in my view, should be
   so understood).

10 The availability of this indirect reading of the normative relevance of
   virtuous character traits already casts severe doubt on the claim that an
   ideal of virtuous character has normative purchase *only if* the average
   human being has a fighting chance of acquiring the relevant character
   trait.

11 Interpreted without restriction, this condition will compel acceptance
   of the unity of the virtues. I reject that conclusion but do not have the
   space to discuss the issue here. See Sreenivasan (2009).

12 In my 2002 (pp. 65–66), I discuss some preliminary evidence suggesting that adherence to this condition would lead to higher cross-situational consistency correlations than Hartshorne and May found. For a much fuller argument to this effect, see now Snow 2010.

13 Merritt *et al.* (2010) do not even mention it.

14 The distinction is not exhaustive, so there are some situationist experiments that are neither iterated trials nor one-time performances (e.g. the Zimbardo prison experiment). My remarks here are confined to one-time performance experiments.

15 Since consistency is necessary but not sufficient for virtue, the question in the text is simpler than asking whether Milgram's subjects lacked the *virtue* of compassion. For discussion of some of the complications I am leaving out here, see Sreenivasan (2008).

16 Degrees of trait reliability are different from dimensions of trait reliability. Temporal stability and cross-situational consistency are the "dimensions" of reliability we distinguished at the outset, while degrees of reliability range from 0 to 100.

17 The lower the degree of reliability required of a trait, the greater the number of failures needed to exclude the trait.

18 But, even if you have these temptations, why suppose that *this* is a better explanation than, say, the explanation that the subject lacks the temporally stable trait of compassionate-under-the-instruction-of-experimental-authorities?

19 Since *virtues* have to be both cross-situationally consistent and temporally stable, we really want evidence from iterated trial experiments that instantiate both species at once, just like Hartshorne and May's. See note 15.

20 It counts the second trait-contrary performance in a given eliciting situation as diminishing the person's trait reliability to the same extent as a trait-contrary performance in any *other* trait-relevant eliciting situation.

21 Of course, we could always add a suitable assumption to this effect. But then we could also save a lot of bother and simply assume the conclusion outright.

22 There is no requirement that the argumentative route to the philosophical situationists' bottom line proceed via their official interim conclusion, although that is how they appear to argue their case. A perfectly good alternative strategy is exhibited by Miller (2009, 2010), for example. In an illuminating discussion of one-time performance experiments about helping behavior, he argues that the experimental record is actually consistent with many people's having *some* cross-situationally consistent helping trait. Yet, at the same time, he argues that, of the

people with some such trait, few actually have the *virtue* of compassion because, in many cases, the cross-situationally consistent helping trait fails a different necessary condition for virtue. (Many of these people characteristically help for the wrong reason, for instance, though there is an interesting variety in their wrong reasons). I do not discuss Miller's argument in the text because he explicitly concedes the scope objection.

23   With *neither* kind of similarity, it may help to spell out, is the operation of the situational influence itself any part of what makes numerically distinct situations similar.

24   Here again, any such argument has to strike a balance between making its claim weak enough to be plausibly grounded in the data and yet strong enough to contradict some actual tenet of virtue ethics.