© Springer 2006

JASON KAWALL

## ON THE MORAL EPISTEMOLOGY OF IDEAL
## OBSERVER THEORIES

ABSTRACT. In this paper I attempt to defuse a set of epistemic worries commonly raised against ideal observer theories. The worries arise because of the omniscience often attributed to ideal observers – how can we, as finite humans, ever have access to the moral judgements or reactions of omniscient beings? I argue that many of the same concerns arise with respect to other moral theories (and that these concerns do not in fact reveal genuine flaws in any of these theories), and further, that we can and often do have knowledge of the reactions of ideal observers (according to standard, prominent theories in the domain of epistemology).

### I

In what follows I attempt to defuse a set of worries, all grounded in epistemic concerns, commonly raised against ideal observer theories. I will focus on ideal observer theories in ethics that characterize moral properties (of actions, states of affairs, and so on) in terms of the judgements or reactions of ideal observers.[1] For example: an action is morally right if and only if an ideal observer would approve of the action (in some particular way). Here moral rightness is constituted by the attitudes of an ideal observer. That is, it is not merely that ideal observers are good epistemic agents who can identify actions that possess a prior property of moral rightness; rather, actions are right precisely because of the pro-attitudes that an ideal observer would have towards them.

Such theories have been attributed to (and endorsed) by many philosophers. For example, both David Hume and Adam Smith have been treated as

---

[1] Ideal observer accounts of other normative properties – in particular, aesthetic properties – are quite common, but will not be discussed here. See, for example, Goldman, (1995) and Taliaferro (1990). Note that the arguments presented in this paper would apply, *mutatis mutandis*, to these other domains.

ideal observer theorists.[2] Roderick Firth presents perhaps the best-known explicit defence of an ideal observer theory in his important "Ethical Absolutism and the Ideal Observer" (Firth, 1952). Richard Brandt embraced the approach in his early work, and the ideal observer also plays a prominent role in the work of R.M. Hare (in the form of the 'archangel').[3] Versions of the view have been endorsed in recent years by such philosophers as Charles Taliaferro, and Thomas Carson.[4] More broadly, ideal observer theories can be understood as a particular form of 'moral-response dependence' theory, a general approach that has received much recent attention.[5]

Why concern ourselves with ideal observer theories? Richard Brandt, in discussing Firth's proposal lists several merits, including:

> (1) that this theory [Firth's] enables us to regard as really relevant to ethics all the facts which on reflection we take to be relevant; (2) that is enables us to explain the heterogeneousness of the actions we regard as right or wrong; (3) that it explains how ethical disagreement is possible even when there is agreement about the nature of the act being appraised; (4) that it explains why our feelings and attitudes – and especially our sympathies – are (and properly are) engaged in ethical reflection, and why moral philosophers have thought that moral experience is distinctively a union of cognition and emotion; (5) that it enables us to hold that moral opinions are subject to objective criticism and are correct or incorrect; that (6) it explains why we value the advice of knowledgeable, impartial, and consistent persons at times of moral decision, and why we reject previous moral opinions of our own which we think reflect self-interest, inconsistency, or lack of information; and that (7) it enjoys advantages over the emotive theory such as the capacity to give a satisfactory analysis of "ethical relevance", and the ability to explain why ethical judgements do not always correspond with favourable or unfavourable attitudes on the part of the judge (Brandt, 1955a, 407).

Broadly then, ideal observer theories at once provide for both cognitivism (that we have moral beliefs, and that these can be correct or incorrect) and an important place for emotion at the core of ethics. A certain objectivity is achieved (through the appeal to full-information or omniscience, and consistency on the part of the ideal observers), but without appeal to 'queer' realist moral facts. Such theories thus capture a range of features that many deem attractive in a metaethical position, and as such they seem worthy of our consideration.

---

[2] See Hume (2000), and Smith (2002). Note that it is rather controversial whether Hume and Smith are in fact best understood as ideal observer theorists. Still, at the very least, they are clear precursors for such.

[3] See Brandt (1957), and Hare (1981).

[4] Taliaferro (1983), and Taliaferro (1997), esp. chapter 7; Carson (1984), and Carson (2000), esp. chapter 8 [here Carson endorses a divine-preference theory in those worlds in which a suitable God exists; otherwise, he suggests a form of ideal observer theory].

[5] See, for example, Johnston (1989), McDowell (1985), Vallentyne (1996), and Wiggins (1987).

Ideal observer theories vary in what traits they attribute to the ideal observers, and in the exact structure of their accounts of normative properties. Still, a common requirement is that an ideal observer be at least fully-informed about any case being judged.[6] More strongly, many ideal observer theorists require that ideal observers be omniscient.[7] For the present discussion, we will focus on versions of the ideal observer (henceforth 'IO') theory that embrace the latter, more demanding omniscience requirement.

The epistemic worries now become obvious. How are we, as mere humans, ever to know what is morally right on an ideal observer theory? We are not omniscient; nor do we possess unlimited reasoning abilities. The perspective of an ideal observer seems distant and removed at best. Feminist philosopher Donna Haraway dismisses appeals to such perspectives as "the god-trick of seeing everything from nowhere" (Haraway, 1991, p. 189). Linda Zagzebski writes that an

> important drawback of a nonactual IO is that we can surmise his response only to the extent that we think that we ourselves approach the standpoint of the IO, at least in imagination. Sometimes we think we can do that, at least temporarily, and so we treat Kant's claim that a happy person without a good will gives no pleasure to an impartial spectator as a discussable item in classrooms and at conferences. But IO theorists typically include more than impartiality in their list of the attributes of the IO, and it can be very difficult to imagine how a being with such attributes would respond. [...] even if impartiality is something that anyone can adopt by effort alone, it is considerably harder to imagine being omniscient and omnipercipient, much less to imagine what our responses would be if we had those attributes (Zagzebski, 2004, pp. 354–5).

If such criticisms are correct, we have reason to reject IO theories as inadequate, providing us only with unattainable, unknowable standards.

Geoffrey Sayre-McCord articulates several of these epistemic worries in the course of a discussion of Hume's ethics:

> Although stable, and presumably univocal in its deliverances, that point of view [that of an Ideal Observer] is not sufficiently accessible. *We* have neither the psychological equipment nor the knowledge required. Our estimates of the Ideal Observer's view of the effects of someone's character will differ in exactly the way our judgements of actual effects will differ. As a result, an Ideal Observer sets an inappropriate standard, not simply because we cannot take up her position ourselves (though we cannot), but because we cannot begin to anticipate what her reactions might be. Ignorant as we all inevitably are of the actual, subtle, and long-term effects of each person's character on everyone who might be affected, even earnest attempts by all to determine how an Ideal Observer would respond would leave us without a common standard around which to coordinate our actions and evaluations (Sayre-McCord, 1994, p. 218).[8]

---

[6] This is the approach of Carson (1984). See also Richard Brandt (1955a, 1955b).

[7] Examples include Firth (1952, 1955). See also Taliaferro (1997), and Carson (2000).

[8] Note that Sayre-McCord presents these as concerns that shape Hume's moral theory;

We can isolate the following specific worries in this passage: (i) We (mere humans) cannot achieve the standpoint of an ideal observer (a concern also raised by Zagzebski in the passage quoted in the previous paragraph). (ii) We cannot even anticipate the reactions of an IO (another worry also raised by Zagzebski). (iii) IOs will make use of long-term, subtle effects of characters, etc. that are unavailable to us. (iv) Because of these factors, the viewpoint of an IO is not a practical standard that can be of use to humans in their normal affairs.


## II


Before addressing these specific concerns, we can begin with some more general points. First, it is worth noting that the sort of epistemic difficulties raised by Sayre-McCord are not unique to the IO approach. Utilitarians will, of course, have difficulty in determining the probable effects of actions in complex situations. Kantians will have difficulty in determining which maxims can serve as universal laws, how to treat humanity as ends in cases where interests conflict, and so on. Thus, we do not have a particular problem for ideal observer theories.

Second, and more importantly, on any plausible theory of morality there will be difficult, complex cases – cases where we as finite humans cannot be certain that we have come to the right answer. But this should not be surprising, and hardly counts as a flaw in a theory. Morality can be complex, and it would be foolish to think that we humans must always have access to its dictates. As in any other field, we can make mistakes. That we as humans cannot always arrive with certainty at the right action in complex cases is simply a reflection of our limited cognitive and emotional capacities.

Still, the epistemic concerns can be pressed in slightly altered forms. The general worry now is that on an IO account of morality, we as limited humans will be ignorant of what is truly right or wrong. We are unable to attain the standpoints of IOs, so that the realm of morality (as it were) will be forever closed to us. Perhaps the attitudes of ideal observers determine legitimate standards of rightness and wrongness, but standards that are beyond our grasp, and of no use to us in daily life. Further, because of this, we will be unable to settle moral disagreements. We can now turn to the objections attributed to Hume by Sayre-McCord, understood in this light.

---

they are not objections raised specifically by Sayre-McCord himself against IO theories.

## III

(a) With respect to the first objection, that we cannot achieve the standpoint of an ideal observer, note that at most we would simply need to predict the reactions or approvals of IOs – and this does not necessarily require taking up the same perspective. Compare: if we want to predict the reactions of a bat to a variety of circumstances we need not take up the precise perspective of a bat; we do not ourselves require powers of echolocation or what-have-you. Rather, we just need good methods of figuring out what their reactions will be. We do not require absolute certainty that we are correct and justified in our moral judgements (which *might* require being able to take on an IO perspective) – we just need to be reliably correct. So while Sayre-McCord and Zagzebski might be correct that we ourselves cannot achieve the standpoint of an IO, this does not yet pose a significant problem as we do not need to achieve an IO's standpoint to make reasonable, epistemically justified moral judgements.

Moreover, the ideal observer theorist is in no way committed to the claim that we must attempt to achieve the perspective of an IO in our moral deliberations. Cynthia Stark brings this out forcefully:

> Nowhere does Firth recommend that people try to emulate the ideal observer when they are reasoning in particular situations. He simply maintains that moral principles are true just in case they would be ratified by an ideal, and hence impartial, observer (Stark, 1997, pp. 481–2).[9]

Many of the epistemic criticisms of IO theories appear to rest on the assumption that IO theorists intend their position as a decision-procedure. That is, it is assumed that IO theorists require us to try to attain or imagine achieving the position of IOs as we consider particular cases. But the IO theorist is providing an account of the nature of normative properties (grounded in the attitudes of IOs); no particular or exclusive decision-procedure is thereby entailed.

(b) Which leads to the second worry (raised by both Sayre-McCord and Zagzebski) – can we even anticipate the reactions of an IO? It seems that with a good base of knowledge and a virtuous character we should be quite capable of making reliable judgements about the reactions of an

---

[9] Stark's paper provides powerful arguments showing that many objections to deontological, contractualist, and IO theories (and their understandings of impartiality) fail to adequately recognize (i) that such theories concern the nature of justification of moral standards, and need not be intended as decision-procedures, and (ii) that such theories can readily allow partial concerns to be relevant to moral rightness (etc.) from an impartial point of view. That is, for example, it seems entirely possible that an impartial, omniscient ideal observer could approve of actions where an agent favours her loved ones.

IO. There is not a devastating epistemic gap here. In most situations we will be adequately informed and of sufficiently good character (or have exemplars to whom we can turn) to reliably anticipate the reactions of ideal observers. This is obviously relevant also to the first worry – while we may lack omniscience, we can still at least approach the standpoint of an IO as we gain relevant knowledge about a given situation, become aware of possible biases that may affect our judgements, and so on. And as we do so, our judgements will come to mirror those of IOs. Of course in some complex situations, we will lack crucial pieces of knowledge, and our judgements will be flawed. But in a broad range of cases the information relevant to appraising the case will be humanly attainable.

Notice further that the IO theorist can readily make use of methods (for moral deliberation) proposed by those who believe IO theories to be flawed. For example, Margaret Urban Walker writes that

> These [IOs, disinterested judges, etc.] are images of transcendence or encapsulation, and their prevalence suggests that an account of morally adequate attention lies not in close perusal of the many talents and techniques that ordinary folk, or persons with specially refined or schooled skills of interpersonal (and self-) observation, make use of to discover "what it is like to be those people in that situation" [. . .] We do not often in articles on moral philosophy see the moral agent at deliberation imagined as a close friend, loving parent, concerned teacher, or perceptive advisor, much less as a gifted counselor, seasoned psychoanalyst, shrewd sociological observer, or trained anthropological field worker. Yet all such individuals possess special capacities and opportunities for gleaning recondite information of just the right sorts in some situations where human interests and perceptions are paramount (Walker, 1991, p. 763).

Similarly, Marilyn Friedman holds that

> If these methods [such as appeal to IOs] for representing impartial normative thought are to provide us with genuine substantive insights into matters of morality or politics, then they must outline methods of reflection that are within the capacities of human beings to adopt (Friedman, 1993, p. 19).

Friedman assumes that IO theorists require us to simply try to attain the standpoint of an IO. In place of such a 'method', she suggests that

> As for methods of eliminating recognizable biases from critical moral thinking, foremost emphasis must go to interpersonal dialogue. For good psychological reasons, each person's unaided thinking cannot be trusted to discern its own biases. One's own thinking – explicit and implicit, avowed and tacit – is not fully transparent to oneself. One's covert racist or anti-Semitic bias, or hostility toward the aged or the disabled, may well be noticeable to others even when invisible to oneself (Friedman, 1993, p. 32).

Walker and Friedman suggest that appealing to ideal observers encourages (and perhaps even requires) us to attempt to attain or imagine a transcendent

view of the world, while ignoring actual embodied, effective methods of overcoming bias, gaining relevant knowledge and so forth. But again: the IO theorist need not hold that we must (or even should) try to directly attain the epistemic standpoint of an IO. Rather, if the methods cited by Walker and Friedman are effective for us in gaining information, exposing bias, and so on (as seems highly plausible), then we ought to use them. They allow us to make informed, grounded judgements, ones that will capture those of IOs (even if we arrive at them via different methods). Put otherwise: why should an IO theorist discourage or deny the use of such effective epistemic instruments?

To this point it seems that we've secured the possibility of moral knowledge on IO theories if we embrace some form of externalist, reliabilist moral epistemology. It is beyond the scope of the present paper to argue for any particular theory of this general kind. But note that – crucially – an IO theorist can appeal to well-regarded epistemological theories to ground the possibility of moral knowledge. There is no need to appeal to an obscure or *ad hoc* theory, ungrounded intuitions, or so on.

Note further that even on an IO theory we can obtain reflective moral knowledge. We do not simply have blind but reliable intuitions about the reactions of ideal observers; we are not mere moral thermometers. Rather, we can in most cases justify our reactions, and justify our beliefs about the reactions of IOs. We can, for example, explain that a given case involves torture, and that this will produce a negative attitude in an ideal observer, just as it does in us. We can explain how our reaction coheres with our reactions in other cases which involve causing suffering, and so on. As such, we will generally be able to achieve reflective moral knowledge, beyond mere 'animal' moral knowledge.[10]

The objection might be pressed – "But how can we be certain that we've correctly picked out the reactions of an IO in any given case? How can we be sure that we know?" So we might still worry that there is an epistemic gap when we attempt to discern the attitudes and reactions of IOs.

Still, the objection now makes demands that have been largely rejected by epistemologists, and not only in the domain of moral knowledge. On one reading, the objection is demanding certainty. To *really* know we must obtain certainty, and we cannot achieve this when contemplating IOs and their reactions. But of course the majority of epistemologists have abandoned certainty as a requirement for knowledge, and instead embraced various forms of falliblism. In the case of IO theories, we are not guaranteed to be correct in our appraisals of IO attitudes, but this is quite compatible with any falliblist epistemology. And again, we have good reason to believe that

---

[10] I use the terms rather loosely here, but roughly follow the distinction drawn by Ernest Sosa. See Sosa (1991a, 1991b).

in standard cases we will be quite reliable in our assessments, as we gain the relevant knowledge, take biases into account, and so on.

The objection could also be read as now demanding that we know that we know the reactions of ideal observers, and charging that we cannot attain this ("But how do we know that we know the reactions of the ideal observers?"). The objection thus quietly assumes a 'KK' principle, such that knowing requires knowing that one knows (Kp→KKp). But such 'KK' principles face important difficulties, and again have been largely abandoned by epistemologists.[11] Perhaps the most troubling difficulty for such principles is that they lead to infinite regresses: To know that p, I must know that I know that p (Kp→KKp). But then to know that I know that p, I must know that I know that I know that p (KKp→KKKp). And so on. We see that the objection, when pressed in these further forms make demands of ideal observer theories that have been rejected by most epistemologists. As such, we can rightfully put the objection aside.[12]

(c) Sayre-McCord next worries that an IO would make use of extremely obscure bits of knowledge in assessing an action, and that this makes it impossible for humans to achieve the viewpoint of an IO, or even predict an IO's reactions. Walker expresses a similar concern, "God sees everything, but how should we imagine the salience and priority for God of what God sees from God's point of view, and what has this to do with (our) morality?"(Walker, 1993, p. 764). Here we can consider two different kinds of 'distant' knowledge that might be at stake.

First, we might worry that an IO's reactions would be influenced by factors or pieces of knowledge that would be entirely unexpected by us. For example, perhaps an IO's reactions to a case in which a human person is contemplating telling a white lie would vary depending on the weather conditions that then obtain on the fifth planet of Alpha Centauri. This would be quite unexpected, and would likely make our appraisals of an IO's attitudes highly unreliable in this sort of case.

Still, while this is possible, it hardly seems likely. We can safely classify scenarios of this kind as akin to the evil demon cases typically considered by epistemologists. The suggestion that IOs would treat such bizarre pieces of information as relevant is unmotivated; surely we deserve at least some explanation of why IOs would react in such ways, if we are to have any reason to take the possibility at all seriously. And again, an appeal to a

---

[11] See, for example Nozick (1981), p. 245ff, and Williamson (2000), ch. 5.

[12] I should perhaps add that I am not arguing that the mere fact that most contemporary analytic epistemologists reject demands for certainty, etc. shows that such demands *must* be mistaken. Rather, I wish only to show that the IO theorist is at least in good company in rejecting such demands, and that the burden of proof rests with those who would claim that these demands must be met.

reliabilist, externalist epistemology can help us to see our way clear of the apparent problem. What matters is that we are in fact reliable in appraising the attitudes of IOs. We cannot guarantee with absolute certainty that IOs do not appeal to such obscure facts, but such a demand for certainty is rightfully rejected (and again – not only in the domain of moral epistemology).

While defending ideal observer theories in general, we can briefly consider the following specific formulation:

> An action is right for an agent in a given set of circumstances iff an unimpaired, omniscient virtuous observer would deem the action to be right (where this is a certain form of approval).[13]

Of course, such a proposal would require an independent account of the virtues. But notice that on this account we would have still greater reason to reject claims that ideal observers might treat intuitively bizarre information as relevant to moral assessment. After all, compassionate, honest, benevolent, and otherwise virtuous humans do not treat such sorts of information as relevant; so again, why think that this would be otherwise with virtuous ideal observers?[14]

Such appeals to actual moral exemplars can be extremely valuable as we attempt to determine the attitudes of ideal observers. In a recent book Zagzebski provides an insightful account of the importance of exemplary persons for moral theory and practice. In particular, she proposes

> a theory of the genealogy of morals according to which we learn moral concepts and acquire the ability to make moral judgments from experience. That includes the experience of imitating the way in which paradigmatically good persons form concepts and make judgments. Emotions are ways of affectively perceiving the world around us that have conceptual constituents of a distinctive kind. We have emotions automatically, but we learn them in part by imitation, and they change under the influence of the emotions of admired others (Zagzebski, 2004, p. 51). Most exemplars are recognized as exemplars, so they are persons consulted for advice by others, and to the extent that it can be confirmed by hindsight, their advice turns out to be correct. They are stable and reliable. They usually have a lot of general knowledge. They have peace of mind. In typical circumstances, they are happy. But they are also prepared to face tragedy, and when it enters their life, they are good at handling it (Zagzebski, 2004, p. 56).

---

[13] For further discussion, see Kawall (2002).

[14] Notice also that this particular proposal would provide important resources to respond to those, like Walker, who maintain that ideal observers might be too detached or isolated to provide proper moral standards. Such virtues as compassion and benevolence will ground the attitudes and responses of the IOs, encouraging a sensitivity to the subtle particularities of given cases. For example, such ideal observers would readily recognize the importance of the friendships, roles, and so on of the particular agents involved.

With respect to moral exemplars, (1) we can learn about their actions and commitments, and use this knowledge as a guide to moral rightness, (2) we can learn of their emotions and attitudes towards various scenarios (both actual and hypothetical), and use this knowledge as a guide to appropriate reactions to such cases, (3) we can seek moral advice and guidance from them (that is, they can themselves attempt to articulate their moral views, and explicitly provide us with advice), and (4) we can imitate them, thus shaping ourselves so that our reactions come to match those of the exemplars; in turn, we should expect these reactions to match those of virtuous IOs.

Let us consider Dorothy Day as a moral exemplar. She was an intelligent and knowledgeable woman, with particularly extensive and intimate knowledge of the working poor, and the unemployed. She co-founded the Catholic Workers movement, and established a 'House of Hospitality' in 1933 in New York City to help the homeless and others struggling through the Great Depression – there are now dozens of such houses across the United States (and in other countries). Throughout her life she was active in a wide range of social causes, especially on behalf of those living in poverty.

We can learn much from the case of Day, and other moral exemplars. To begin with, and to return to the worry that initiated this discussion, there is no evidence that Day (or any other exemplars) treated bizarre bits of information as morally relevant; she does not appear to have taken into account the motion of comets in some distant galaxy as she deliberated. This gives us good grounds to expect that virtuous IOs would also treat such information as irrelevant to moral decision-making. Furthermore, she undertook many commitments, treating the suffering of others as significant, and fighting against what she deemed to be important injustices. We thus have good grounds to hold that such projects would be approved of by virtuous IOs (after all, a well-informed and virtuous person treated these projects as morally significant; and other well-informed exemplars have undertaken similar projects). The burden of proof would surely be on those who would claim that – somehow – with additional information or more virtuous characters, Day and other moral exemplars would no longer approve of such commitments, or would begin to treat bizarre, intuitively irrelevant facts as somehow relevant to moral deliberation. Of course as finite humans, moral exemplars can make mistakes (due to lack of information, or lapses of virtue) – their approvals are not an infallible guide to the attitudes of virtuous IOs – but this is quite compatible with their attitudes serving as highly reliable guides to the judgements of IOs.[15]

---

[15] Note that in the case of exemplars whom we do not know personally, we primarily gain knowledge about the kinds of actions and commitments that virtuous IOs would approve of.

We can now turn to the second kind of 'distant' knowledge that might influence ideal observers. The knowledge at stake would be of such things as unexpected consequences of actions that humans would be incapable of predicting, facts about the true inner motivations of a given agent – motivations that would be inaccessible to humans (including perhaps the agent herself), and so on. Here we might not have access to the IO viewpoint or reactions, but this seems to reflect human limitations, not a flaw in the IO theory. Some knowledge will be unavailable to most humans, but crucial to determining the true moral status of an action (consider perhaps the actions of an extremely good liar). Still, while there will be cases of this kind, we have no reason to think that this will typically be the case. In most day-to-day moral quandaries we can expect enough relevant information to be available to us to allow us to form epistemically justified beliefs about the reactions of IOs to such cases.

(d) With this in hand, we can consider the broader worry that an IO standard will not be a practically viable standard for humans to live by. And we can enter the issue by considering the alternative 'general point of view' that Sayre-McCord attributes to Hume. Sayre-McCord argues that the general point of view is attainable by humans, and can therefore serve as a practical standard for making moral judgements, unlike the unattainable position of an ideal observer:

> When it comes to morality, Hume holds that virtually all of us are qualified to judge, so long as we take into account only our sympathetic responses to people's characters, control for distortions of perspective, and focus on the tendencies rather than the actual effects of the characters judged on those in the "narrow circle". In taking up that point of view [the general point of view], we need know neither all the actual effects of the person's character nor the usual effects on all (Sayre-McCord, 1994, p. 212).

The general point of view (GPOV) is thus a viable practical standard because, according to Sayre-McCord, it is a standpoint that is accessible to almost all normal adult humans. Further, we can settle moral disagreements because we can appeal to this attainable point of view.

But is this general point of view any easier to attain than that of an ideal observer? Notice first that we would surely require knowledge of an agent's character and motives which we, as mere humans, sometimes

---

That is, we learn about the actions and commitments of well-informed, virtuous persons, which serves as evidence that virtuous IOs would approve of such actions and commitments. It is less feasible for us to imitate their general ways of being and emotions as we lack immediate, day-to-day encounters with them. For the latter, personal acquaintance with a moral exemplar would allow for more effective imitation. And again, as we model our reactions, behaviour, and emotions on those of moral exemplars, our attitudes and emotions will come to match those of these exemplars; in turn, we should expect that our attitudes and reactions would match those of virtuous IOs.

do not have. If we appeal only to the knowledge of an agent's motives (and the effects they tend to produce on the agent's 'narrow circle') that would be generally available, we would be subject to deception. Apparently benevolent actions may be driven by selfish motives, and so on. We also have limited access to an agent's underlying character and its real (as opposed to merely apparent) tendencies in affecting others. If so, while we may have a more accessible viewpoint in the GPOV, it is one which would seem to carry little normative weight. It would be too prone to error. On the other hand, if achieving the general point of view requires knowledge of the actual motives and underlying character of an agent, we would require faculties and knowledge beyond those of normal humans. If so, the GPOV is not interestingly more accessible to normal humans than the IO point of view.[16]

Second, Sayre-McCord's Hume would have us focus on the effects an agent's character would have on those in her narrow circle. How do we determine what this narrow circle is? Hume notes that the narrow circle of a politician may extend across an entire nation. With others, the narrow circle may include only the person's family and close friends. There are, in fact, several questions here: (a) what determines the range of the narrow circle (perhaps the judgements of an ideal observer?), (b) how do we as humans come to know the correct range (and would this require knowledge beyond that commonly available to humans?), (c) how much must we know of the character of those in the narrow circle (are they jealous, vindictive, etc.?), and (d) if the narrow circle is extended (as in the case of a politician), would we not require mental abilities beyond the range of normal humans in order to ascertain the impact of an agent's character on the entire circle? It seems that the general point of view would be just as difficult to achieve as that of an ideal observer (though again, this inability to achieve a standpoint itself is not necessarily troubling from an epistemic point of view, so long as we can reliably form beliefs about the attitudes which would be taken by an observer with the given standpoint). And to the extent that the GPOV isn't as difficult to achieve, it becomes inadequate as important, relevant facts are ignored or excluded from consideration.

Third, suppose we have two agents who are attempting to assess an action, and thus attempting to take up the GPOV. One of the agents manages to take up the GPOV, while the other does not. Sayre McCord claims that the GPOV provides us with an attainable point of view. Even if it is attainable (which we have seen is rather dubious), it will not solve the problem of massive disagreement. Take our two agents – both will presumably think they have taken up the GPOV, even if only one of them actually has done

---

[16] The GPOV might be more accessible to humans in the same way that running at 100 km/h is more accessible to us than running at 150 km/h. . .

so. Surely they will still disagree about who has achieved the GPOV. And if so, Sayre McCord's proposed move to the GPOV will not eliminate disagreement in the way that he suggests.[17] Basically put – humans could disagree about who has truly achieved the point of view of an ideal observer, but it seems that they would also disagree about who has truly taken on the GPOV. There is no advantage in shifting to the GPOV.

We can now draw out a broader lesson. People will be quite capable of disagreeing about the application of any plausible moral theory. That is, people can disagree about the proper application of the categorical imperative, what will in fact maximize utility, and so on. This is a product of human fallibility and is not obviously a problem with any of these theories. Furthermore, in the case of IO theories, we can take important steps towards resolving disagreements – the most obvious of these is to ensure that all disputants have access to as much (and the same) relevant information as possible, while accounting as far as possible for potential biases and prejudices. And this can be pursued via, among other methods, the interpersonal dialogue and attention to particularity emphasized by Friedman and Walker, and the appeals to moral exemplars articulated by Zagzebski.

IV

We can conclude with two final observations. First, we might suspect that many people place great weight on the epistemic objections to IO theories because they worry that humans will be unfairly held to moral standards to which they lack access. It might appear that on ideal observer theories people could be blamed for not performing morally right actions, but in cases where the reactions of an omniscient IO were simply inaccessible to normal human agents. And of course, this strikes us as unjust.

We have already seen how we can have reliable access to such reactions in typical cases, even while we might be incapable of taking up an ideal observer's viewpoint ourselves. But what of those cases where we lack crucial information, such that we lack reliable access to IO reactions? Here we can distinguish between morally justified and morally right actions. Roughly, morally justified actions are those which would be appropriate or permissible for a moral agent, given her (non-culpable) epistemic and other limitations (in given circumstances).[18]

---

[17] It should perhaps again be stressed that Sayre-McCord does not explicitly endorse these arguments himself; he instead presents them as factors that influenced Hume's moral thinking.

[18] Similar accounts of morally justified actions can be found, for example, in Goldman, A.I. (1980), Feldman (1988), Montmarquet (1993), and Dancy (1993, p. 113).

Ideal observers can distinguish between what is morally right and what a given agent, given her actual situation and limited knowledge is morally justified in doing. Ideal observers, being omniscient, will of course recognize that we sometimes lack (and cannot obtain) important information about situations that we confront – and can take such limitations into account as they approve of actions as morally justified for us. Such justified actions might not always be right actions, but we need not see such agents as blameworthy, so long as they are deemed justified in their actions by an IO. Thus we might index blameworthiness or praiseworthiness to moral justification; praise and blame need not be directly tied to strict moral rightness.

Note further that this is a common distinction, and not an *ad hoc* manoeuvre on behalf of IO theories. In ordinary moral thought, in cases where a person is appropriately careful and acts in good faith but due to strange circumstances fails to perform what is in fact morally right, we typically do not hold such an agent to be blameworthy, given her care and effort. Her actions are morally justified, even if not right. And there are common examples reflecting the other side of this coin; cases in which an agent happens to perform a morally right action, but where this action was not justified for the agent, and the agent is not deserving of moral praise.[19]

The second observation is the following: we do not need to explicitly determine what IOs would approve of in order to perform morally right (or justified) actions. For example, imagine a case in which a parent rushes into a house to save his child from a fire. The father acts out of love for his child, and does not pause to contemplate the reactions of ideal observers. IOs could still clearly approve of the father's action as right, even while he did not pause to determine explicitly the morally right action; indeed, the approval might be even greater insofar as the father did not have 'one thought too many'. Even if there are some cases in which we lack reliable access to the reactions of IOs, we can still perform morally right (and/or justified) actions. Our actions need only to be suitably approved of by ideal observers; as noted in our discussion of Sayre-McCord's first objection, IO theorists are in no way committed to the claim that we must attempt to achieve the standpoint of an IO in our moral deliberations. Thus, even in worst case scenarios, where we lack epistemic access to the reactions of IOs, this does not preclude our performing right actions. And as such, we

---

[19] Consider the actions of a lunatic who adds fluoride to a village's water supply in attempt to poison them. As it turns out, he adds an ideal amount, providing villagers with stronger teeth, and inadvertently performing a morally right action. Surely he still deserves no moral praise. Or consider a case in which you promised to meet a friend for lunch, entirely forget about the promise, but quite by chance decide to have lunch at the same restaurant, and so (by chance) happen to meet your friend, as promised.

can see that ideal observer theories are not at any disadvantage (relative to other viable moral theories) in providing accessible, practical moral standards.

## REFERENCES

Brandt, R.B., *Ethical Theory*. Englewood Cliffs: Prentice-Hall, 1957.

Brandt, R.B., The Definition of an 'Ideal Observer' Theory in Ethics, *Philosophy and Phenomenological Research* 15(MR) (1955a), pp. 407–413.

Brandt, R.B., Some Comments on Professor Firth's Reply, *Philosophy and Phenomenological Research* 15(MR) (1955b), pp. 422–423.

Carson, T., *The Status of Morality*. Dordrecht: D. Reidel, 1984.

Carson, T., *Value and the Good Life*. Notre Dame: University of Notre Dame Press, 2000.

Dancy, J., *Moral Reasons*. Oxford: Blackwell, 1993.

Feldman, R., Subjective and Objective Justification in Ethics and Epistemology, *The Monist* 68(JL) (1988), pp. 407–419.

Firth, R., Ethical Absolutism and the Ideal Observer, *Philosophy and Phenomenological Research* 12(MR) (1952), pp. 317–345.

Firth, R., Reply to Professor Brandt, *Philosophy and Phenomenological Research*, 15(MR) (1955), pp. 414–421.

Friedman, M., *What Are Friends For? Feminist Perspectives on Personal Relationships and Moral Theory*. Ithaca, NY: Cornell University Press, 1993.

Goldman, A.H., *Aesthetic Value*. Boulder: Westview, 1995.

Goldman, A.I., The Internalist Conception of Justification, in P. French, T. Uehling, and H. Wettstein (eds.), *Studies in Epistemology*, Midwest Studies in Philosophy, v. 5. Minneapolis: University of Minnesota Press, 1980, pp. 27–51.

Haraway, D., Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective, essay 9 in *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge, 1991.

Hare, R.M., *Moral Thinking*. New York: Oxford University Press, 1981.

Hume, D., *A Treatise of Human Nature*. New York, Oxford University Press, 2000.

Johnston, M., Dispositional Theories of Value, *Proceedings of the Aristotelian Society* 63(Supp.) (1989), pp. 139–174.

Kawall, J., Virtue Theory and Ideal Observers, *Philosophical Studies* 109(3) (2002), pp. 197–222.

McDowell, J., Values and Secondary Qualities, in T. Honderich (ed.), *Morality and Objectivity*. London: Routledge and Kegan Paul, 1985, pp. 110–129.

Montmarquet, J., *Epistemic Virtue and Doxastic Responsibility*. Lanham, MD: Rowman and Littlefield, 1993.

Nozick, R., *Philosophical Explanations*. Cambridge, MA: Harvard University Press, 1981.

Sayre-McCord, G., On Why Hume's 'General Point of View' Isn't Ideal – And Shouldn't Be, in E.F. Paul, F. Miller, and J. Paul (eds.), *Cultural Pluralism and Moral Knowledge*. New York: Cambridge University Press, 1994.

Smith, A.., *The Theory of Moral Sentiments*. New York: Cambridge University Press, 2002.

Sosa, E., Knowledge and Intellectual Virtue, Essay 13 in *Knowledge in Perspective*. New York: Cambridge University Press, 1991a.

Sosa, E., Intellectual Virtue in Perspective, Essay 16 in *Knowledge in Perspective*. New York: Cambridge University Press, 1991b.

Stark, C., Decision Procedures, Standards of Rightness, and Impartiality, *Noûs* 31(4) (1997), pp. 478–495.

Taliaferro, C., The Divine Command Theory of Ethics and the Ideal Observer, *Sophia (Aust.)* 22(JL) (1983), pp. 3–8.

Taliaferro, C., The Ideal Aesthetic Observer Revisited, *British Journal of Aesthetics* 30(1) (1990), pp. 1–13.

Taliaferro, C., *Contemporary Philosophy of Religion*. Oxford: Blackwell, 1997.

Vallentyne, P., Response-Dependence, Rigidification, and Objectivity, *Erkenntnis* 44(1) (1996), pp. 101–112.

Walker, M.U., Partial Consideration, *Ethics* 101(4) (1991), pp. 758–774.

Wiggins, D., A Sensible Subjectivism?, Essay V in *Needs, Values, and Truth*. New York: Oxford University Press, 1987.

Williamson, T., *Knowledge and Its Limits*. New York: Oxford University Press, 2000.

Zagzebski, L.T., *Divine Motivation Theory*. New York: Cambridge University Press, 2004.

*Department of Philosophy and Religion*
*Colgate University*
*13 Oak Dr.*
*Hamilton NY 13346,*
*USA*
*E-mail: jkawall@mail.colgate.edu*