

# MORAL AGENT AND IMPARTIAL SPECTATOR

by

GILBERT HARMAN



The Lindley Lecture  
The University of Kansas  
1986

The E. H. Lindley Memorial Lectureship Fund was established in 1941 in memory of Ernest H. Lindley, Chancellor of the University of Kansas from 1920 to 1939. In February 1941 Mr. Roy Roberts, the chairman of the committee in charge, suggested in the *Graduate Magazine* that

the Chancellor should invite to the University for a lecture or a series of lectures, some outstanding national or world figure to speak on "Values of Living"—just as the late Chancellor proposed to do in his courses "The Human Situation" and "Plan for Living."

In the following June Mr. Roberts circulated a letter on behalf of the Committee, proposing in somewhat broader terms that

The income from this fund should be spent in a quest of social betterment by bringing to the University each year outstanding world leaders for a lecture or series of lectures, yet with a design so broad in its outline that in the years to come, if it is deemed wise, this living memorial could take some more desirable form.

The fund was allowed to accumulate until 1954, when Professor Richard McKeon lectured on "Human Rights and International Relations." The next lecture was given in 1959 by Professor Everett C. Hughes, and has been published by the University of Kansas School of Law as part of his book *Students' Culture and Perspectives: Lectures on Medical and General Education*. The selection of lecturers for the Lindley series has since been delegated to the Department of Philosophy.

# MORAL AGENT AND IMPARTIAL SPECTATOR

by

GILBERT HARMAN  
Professor of Philosophy  
Princeton University

The Lindley Lecture, University of Kansas, April 18, 1986

© Copyright 1986 by Department of Philosophy  
University of Kansas

# Moral Agent and Impartial Spectator

Gilbert Harman

One important type of ethical theory treats moral properties as analogous in certain respects to “secondary qualities” like colors. According to this sort of theory, whether something is right or wrong depends on how impartial spectators would react to it. In the 18th Century, the Scottish philosophers Francis Hutcheson, David Hume, and Adam Smith explored theories of this type.<sup>1</sup> In the 20th Century this sort of theory has sometimes been discussed under the name “ideal observer theory.”<sup>2</sup> Recently, especially in England, there has been renewed interest in this sort of ethical theory and its comparison between moral properties and secondary qualities.<sup>3</sup>

One possible objection to an impartial spectator theory is that it seems to require an overly aesthetic conception of morality to take the primary point of view in ethics to be that of a spectator rather than that of the agent.<sup>4</sup> If the spectator is taken to be primary, then the agent’s aim would seem to be to produce something that will or would please the spectator. But that is just wrong. Such an aim is too “outer directed” to count as a moral motive. Morality is more agent-centered than that. It is much more plausible to take the agent’s point of view as primary. In the first instance morality is a matter of the moral reasons an agent has to act in one way or another, where these reasons derive from the relevant moral rules rather than from a desire to gain the approval of spectators.

A few years ago, I pressed this objection myself when I discussed the ideal observer theory in a textbook of ethics.<sup>5</sup> But I was too hasty. In reading Adam Smith’s *Theory of the Moral Sentiments*, I discovered that Smith explicitly considers this issue and provides a plausible reply to the objection.

Because 20th Century discussions have tended not to consider such “psychological” questions as why agents might be motivated to act in ways that impartial spectators would approve (or, for that matter, why impartial spectators would care about anything), I will in this paper ignore recent discussion and return to the three great versions of the theory that were developed in the 18th Century by Hutcheson, Hume, and Smith, indicating why I think Smith’s version of the theory is superior to the others.

## *Analogy with Secondary Qualities*

Hutcheson and Hume each modeled an impartial spectator theory of moral right and wrong after a normal-perceiver theory of secondary

qualities like color, taste, sound, etc. The theory of color, for example, takes facts about colors to be facts about the reactions of normal perceivers under ideal lighting conditions. For an object to be red is, in this view, for the object to be such that it would look red to normal observers in good light. Similarly, according to Hutcheson and Hume, for an action to be wrong is for the action to be such that it would displease normal observers under conditions ideal for reacting to actions.

This sort of account explains a property by appeal to the psychological experience of an observer. Redness is explained in terms of an observer's visual experience—how things look to an observer. Wrongness is explained in terms of *moral experience*—how things feel to an observer, what attitude (positive or negative) the observer takes.

The relevant experience does not have to be actual. It is the experience an observer *would* have under certain ideal conditions. In yellow light, an object may look orange to observers but actually be red—if it would look red to normal observers in good white light. An action can seem hateful to actual biased or uninformed observers but really be morally right—if the action would be favored by impartial spectators who knew all the facts.

In fact, this distinction between actual and merely hypothetical reactions was not so clear in the 18th Century. Philosophers who adopted this general approach were apt to worry over the question whether a falling tree makes any sound if no one hears it. The correct answer should have been that the tree does make a loud sound, because, if someone had been there, he or she would have heard something. But philosophers did not always see the possibility of this response. Berkeley thought that the only way to allow for the sound of the falling tree was to have the sound heard by God. Indeed, God was needed even for the tree, since without God's perception, not only would there be no sound, but the tree itself would not be there to fall, according to Berkeley. Hume rejected God as a way of saving the falling tree and its sound and concluded that the unperceived tree was a confused fiction we postulate to give order to our experience of the forest.

But, although 18th Century proponents of theories of this sort do not clearly distinguish between actual and hypothetical reactions of spectators, once the distinction is recognized, it is clear that theories of this sort should refer to hypothetical reactions rather than actual reactions of observers.<sup>6</sup>

Even so there are problems.<sup>7</sup> It seems possible that there might be a red object that would turn green if placed in good light. This red object

would look green to normal observers if placed in good light. How can that be, if for the object to be red is for it to be such that it would look red to normal observers if placed in good light? One possible answer is that the object would indeed look red if placed in good light *and* the object were not to change color. (But then what is it in this view for something to change color?)

Similarly, it seems possible that there could be a wrong action that would have been right if only the act had been considered by impartial spectators. Its bad consequences might outweigh its good consequences, but, if it had been considered by impartial spectators as an example, it would have had enough additional good consequences to turn the act into the right thing to do! Then we seem to have a wrong action that would be approved by impartial spectators if they considered it, again violating the claim of this sort of spectator theory. A possible response is to say that an act is wrong if the act would be disapproved of by impartial spectators who considered the act as it actually was, ignoring any features or consequences attaching to the act through its being considered by the spectators. However there are probably other versions of this objection that escape this response.

But let us forget about these (admittedly serious) problems to return to the basic idea behind impartial spectator theories, namely, that the rightness or wrongness of actions is dependent on the actual or hypothetical reactions of impartial spectators in a way that is analogous to the way in which what color something is depends on how it looks or would look to normal observers in good light.

Of course, the analogy can only be partial. One important difference is that an impartial moral spectator does not have to *perceive* the act being judged. It is enough for the “spectator” to be given a sufficiently full description of the act. Indeed, the act itself may well be merely hypothetical. The agent may be considering whether to do it or not. If the act is wrong, let us hope that the agent does not carry it out. In that case, the act does not exist.

The important point of the analogy is that, just as the theory of color takes color judgments to be about the hypothetical reactions of normal perceivers in ideal situations, an impartial spectator theory takes moral judgments to be judgments about the hypothetical reactions of impartial, knowledgeable people. So, in this view, to say that something is morally right is, roughly speaking, to say that impartial spectators are or would be in favor of it. To say that something is wrong is to say that impartial spectators are or would be opposed to it.

### *Key Issues for Impartial Spectator Theories*

I now want to consider the following two issues for impartial spectator theories in ethics. First, why should an agent care about the reactions of spectators? Second, why should an impartial spectator care about what an agent does?

The first issue lies behind the objection to impartial spectator theories that I have already mentioned. According to such a theory, the desire to do what is right is the desire to act in a way that spectators will approve. But that is too "outer directed". Such a desire is precisely not a desire to do something simply because it is right.

Impartial spectator theories might try to avoid this objection by arguing that agents are not motivated directly to do what spectators would approve of. A moral agent's intention is not of the form, "Let me do what would gain an impartial spectator's approval." Rather, the agent's intention has the form, "Let me do *D*," where in fact doing *D* will be something that impartial spectators would approve but that fact is not the agent's reason for doing *D*. Now, an adequate impartial spectator theory cannot treat it as a mere coincidence that moral agents are motivated to act in ways that impartial spectators would approve. So this leaves the problem of explaining how moral agents come to act in ways that impartial spectators would approve of without the agents' having the specific intention to act in that way.

The second issue for an impartial spectator theory concerns why the spectator cares about what the agent does. This issue lies behind the question whether the reason an act is right is that impartial spectators favor it or rather the reason that impartial spectators favor the act is that the act is right. In the dialogue *Euthyphro*, Plato has Socrates ask a similar question of the view that something is good if and only if it is beloved by the gods: Are actions good because they are loved by the gods or are the relevant actions loved by the gods because the actions are good? Socrates raises the question in order to insinuate that the correct answer is the second one, good actions are loved by the gods because the actions are good, whereas the theory he is discussing must argue that the correct answer is the first one, good actions are good because they are loved by the gods.

Both answers cannot be right. It cannot be true both (1) that for an action to be good is for it to be loved by the gods and also (2) that what the gods love about the action is that the action is good. For that would imply that what the gods love about the action is that they love it, which is perverse.

Similarly, it might be argued against an impartial spectator theory that the second option is the correct one: truly impartial spectators favor actions that are right because the actions are right. Impartial spectator theories are necessarily committed to taking the first option here: acts are right because they would be favored by impartial spectators, who must therefore favor these acts for other reasons than that the actions are right. For it would be perverse to suppose that what impartial spectators would favor in these acts is that the acts would be favored by impartial spectators. This leaves the problem of saying what it is that would lead impartial spectators to favor one or another course of action.

To summarize, any impartial spectator theory in ethics needs to say what explains an agent's moral motivation, what explains a spectator's reaction, and why these two things, agent's motivation and spectator's reaction, should be correlated with each other.

### *Hutcheson's Theory*

Francis Hutcheson (1694–1746) appeals to *benevolence* to answer these questions. In his view, people are innately disposed to like other people's being happy and to dislike others' being unhappy. On the one hand, this tendency motivates agents to act so as to make other people happier. On the other hand, it leads spectators to favor such actions and to oppose actions that have the opposite tendency.

In Hutcheson's view an agent acts morally *in order to make other people happier*. The agent is not motivated to act so as to gain the approval of spectators. To be sure, the agent is motivated to act in a way that spectators would approve. Although this is not an intended aspect of the agent's action, it is also no accident, because the same sort of benevolence that leads the agent to act as he or she does also leads the spectator to approve of so acting.

The agent's act is right because impartial spectators would favor the agent's acting in that way. Spectators would not in the first instance favor the act because the act is right but would favor the act because they want people to be happy and the act makes people happy.

This view yields classical utilitarianism. In Hutcheson's words, "that Action is best, which procures the greatest Happiness for the greatest Numbers; and that, worst, which, in like manner, occasions Misery."<sup>8</sup>

This is an elegant theory, but it faces a serious objection. The account of moral motivation is implausible. Benevolence is too weak a motive to be identified with the motive to do what is right. People are

sometimes motivated by benevolence to try to improve the lot of other people, but this motivation is normally quite feeble when compared with ordinary people's aversions to murder, injuring others, stealing, lying, and failing to keep their promises or pay their debts. Generalized benevolence is normally a much weaker motive than self-interest. But the moral motives just mentioned—to avoid killing others, and so forth—are often just as strong as and sometimes stronger than self-interest. Generalized benevolence cannot be the whole story about moral motivation.

### *Hume's Theory*

Impressed by these facts about moral motivation, Hume, like Thomas Hobbes (1588-1679)<sup>9</sup> and Bernard de Mandeville (1670-1733),<sup>10</sup> sees self-interest as an important aspect of the motivation of certain moral acts, namely acts of "justice" such as telling the truth and keeping promises. But, unlike Hobbes and Mandeville and like Hutcheson, Hume argues that benevolence (based on what he calls "sympathy") is another aspect. Self-interest is involved because, if you cannot be trusted to tell the truth, keep your promises or avoid injuring your associates, people will not join up with you in common enterprises and you will lose out in comparison with other people who do tell the truth, keep their promises, and avoid injury to associates. So you have a strong self-interested reason to keep your obligations.

In this view, self-interest leads people to enter into cooperation with others. Over time, cooperation becomes more formalized into a system of cooperation. The overall system of cooperation makes people better off, so benevolent spectators approve of this self-interested motivation. It therefore counts as moral motivation.

Hume allows for benevolent motivation in addition to self-interest. Benevolent feelings in the agent may reinforce self-interested behavior, leading to even stronger motivation. But benevolence by itself is a weak motive. So the motive to be charitable to others is weaker than the motive to keep your obligations. Charity is not in your interest in the way that promise keeping is.

If benevolence plays only a small role in moral motivation, it plays more of a role in explaining the reactions of spectators, in Hume's view. Impartial spectators are (by definition) not personally affected by the agent's act, so self-interest does not favor or oppose the act. Benevolence is therefore the sole source of a spectator's approval or disapproval.

Hume offers an explanation of benevolence in terms of what he calls "sympathy". This contrasts with Hutcheson who simply postulates that God has made us benevolent.

Hume's account of sympathy goes like this. To think of someone else as in pain is to have a painful feeling yourself but located in your image of the other person. To think of another person as pleased is to have a pleasurable feeling located in your image of that person. So, it is pleasant to think that others are happy and unpleasant to think that others are unhappy. That is why impartial spectators prefer agents to do things that tend to make people happier.

For Hume the association of ideas plays an important role in determining what spectators will approve of. When the spectator thinks of the agent acting in a certain way, association of ideas leads the spectator to think of the *typical* effects of such an action. This thought makes the spectator happy to the extent that these envisioned typical effects involve happiness. So, to the extent that Hume's theory is utilitarian, it tends toward rule utilitarianism rather than act utilitarianism.

Roughly speaking, act utilitarianism says that an act is right to the extent that *that very act* makes people happier or less unhappy. Rule utilitarianism says that an act is right to the extent that *acts of that sort* tend to make people happier or less unhappy. Now, any given particular act is of many different sorts of act, so a crucial question for rule utilitarianism is to decide which are the relevant sorts of act for the purposes of moral evaluation. In Hume's version of the theory, this is a psychological question: in thinking about a particular act, a spectator will associate that act with various other acts; this psychological association determines what the relevant sorts of act are for the purposes of moral evaluation.

Hume's theory has little difficulty with the second problem facing an impartial spectator theory, the problem of accounting for the spectator's approval of right actions without supposing the spectator approves of them because they are right. In Hume's view a spectator's approval arises simply from the spectator's sympathetic thought of the pleasures and pains produced by the agent's act and any associated acts, not from any judgment that the act is right. The act is right because it is favored, not favored because it is right.

But Hume's account of moral motivation is still not very plausible. In his view, a moral agent acts largely out of self-interest or out of habits for which there is a self-interested justification. This is an outer-directed motivation—the agent is concerned with the reactions of others because he or she wants them to continue dealing with him or her.

This seems wrong. As Kant objects, a shopkeeper who gives children the correct change because it would be bad for his business if people

were to think he or she cheated children is doing the right thing, but not acting from a moral motive. The shopkeeper's act has no particular moral worth.<sup>11</sup>

Hutcheson identifies the moral motive with benevolence. That seems a motive of the right sort to be a moral motive—it is aimed in the right direction—but it is not a strong enough motive. (Kant famously disagrees, holding that benevolent motivation too is of no moral worth. But here Kant's view is counter-intuitive.) People's moral motivation is much stronger than their benevolence. Hume takes the moral motive to be mainly self-interest, which is a strong enough motive, but a motive that does not seem to be aimed in the right direction.

### *Adam Smith's Theory*

Adam Smith (1723–1790) bases his positive theory on an insightful criticism of Hume's account of sympathy. Smith points out that Hume is wrong to suppose that merely knowing what another person feels is sufficient for sympathy with that person. You might very well not sympathize with another person's feelings if you thought those feelings were inappropriate to the situation. Consider someone who is very upset over a minor scratch on his knee from a fall, for example. You can understand perfectly well how upset the person is without at all sympathizing with the person's extreme agitation.

Smith observes, further, that Hume is also wrong in taking sympathy with another person's pain to be always entirely unpleasant. It is more unpleasant to have to have dealings with someone who is inappropriately unhappy, too upset about something the person ought to treat as a trifle, than it is to have dealings with someone who is upset to the same extent but for an appropriate reason. In the first case, you do not sympathize with the person; in the second case you do sympathize with the person. The other person's pain is easier to take if you can sympathize with that person.

Smith thinks Hume is right about the importance of sympathy in ethics but wrong about what sympathy is and why it is important.

For Smith, the key point is that sympathy is desirable. Not only do spectators want to be able to sympathize with agents, but agents also want the sympathy of spectators. This gives agents a motive to try to have reactions of the sort that spectators can sympathize with.

Furthermore, in Smith's view, "Sympathy . . . does not arise so much from the view of the passion, as from that of the situation which excites it."<sup>12</sup> This gives a spectator a standard for judging the other agent. The spectator imagines him or herself in the circumstances of the

agent and imagines how he or she would react in those circumstances. If the agent's reaction is similar to the reaction the spectator imagines having, the spectator sympathizes with the agent. If the agent's reaction is more extreme than the spectator's imagined reaction, the spectator does not sympathize with the agent.

Spectators approve of reactions they can sympathize with and disapprove of reactions they cannot sympathize with. In Smith's words, "To approve or disapprove . . . of the opinions of others is acknowledged, by every body, to mean no more than to observe their agreement or disagreement with our own. But this is equally the case with regard to our approbation or disapprobation of the sentiments or passions of others."<sup>13</sup>

Smith's approach leads to a very different conception of the content of ethics. Hutcheson's and Hume's theories imply utilitarianism. The aim is to maximize happiness and minimize unhappiness. Smith's theory has no such implication. Spectators do not just add up pleasures and pains.

Smith's normative theory is more stoical than utilitarian. His approach puts considerable importance on self-control. According to Smith, the feelings an ordinary nonideal spectator imagines having in a given situation are rarely as intense as the feelings the spectator or someone else would have in that situation. What an actual spectator imagines an agent feeling is rarely as intense as what the agent actually is feeling. So, the reaction the spectator imagines having in the agent's circumstances tends not to be as extreme as the agent's actual reaction would be *if* that reaction were not influenced by the thought of how the reaction might appear to spectators. Since a person wants sympathy after stubbing a toe, the person will not for long yell and shout and moan, but will try to restrain his or her feelings so as not to appear ridiculous.

Of course, a more knowledgeable impartial spectator would not underestimate the intensity of an agent's pain. But such a spectator would still favor restraint in the agent's reactions, because the reaction that the spectator (correctly) envisions having in the agent's circumstances would be restrained. The spectator would by now have acquired a habit of restraint in reactions as a way of obtaining the sympathy of people who do underestimate the intensity of pains in others.

In Smith's view, the spectator's reactions are heavily influenced by convention. A spectator tries to imagine how he or she would react in the agent's circumstances. But, if the spectator were in the agent's circumstances, he or she would try to modify an otherwise natural reaction so

as to accord with the imagined reactions of other not necessarily ideal spectators. This means that a spectator will be influenced strongly by his or her expectations of how people ordinarily act and react. Conventional ways of acting and reacting serve as evidence about the feelings of other impartial spectators. When a given spectator is imagining how he or she would react, since the spectator will imagine acting in ways that other spectators will sympathize with, what he or she imagines will be skewed in the direction of the conventional reactions. So, Smith's theory is much more conventionalistic than Hutcheson's or even Hume's. Hume takes convention to be important because conventions are useful: people are happier because of what they can accomplish when they adhere to conventions. But for Smith conventions have a more direct effect. The conventions a spectator participates in determine how the spectator will react and so determine what reactions the spectator will sympathize with.

How does Smith handle the problems with which we began? His response to the second problem concerning spectators' reasons for favoring certain actions is the same as the response made by Hutcheson and Hume. The spectators approve or disapprove of actions depending on whether or not they can sympathize with them. Acts are wrong because spectators disapprove of them, i.e. because they do not sympathize with them. It is not that the spectators disapprove of the acts because the acts are wrong.

But what about the first problem concerning the motivation of moral agents? Hutcheson's appeal to benevolence did not work because of the weakness of benevolence. Hume's appeal to self-interest refers to a strong enough motive, but one that is too outer directed if it is involved in the way that Hume thinks it is involved. It may seem that Smith's theory resembles Hume's in this respect. Here again it looks as if the envisioned source of moral motivation is strong enough but too outer directed. Smith seems to be saying that moral motivation is motivation to act so as to obtain the sympathy of spectators. That seems quite wrong.

But in fact, Smith explicitly denies that moral motivation is motivation so as to obtain the sympathy of others. An agent is motivated to be *worthy* of praise.

The love of praise-worthiness is by no means derived altogether from the love of praise. . . .

The love and admiration which we naturally conceive for those whose character and conduct we approve of, necessarily dispose us to desire to become ourselves the objects of the like agreeable sentiments,

and to be as amiable and as admirable as those whom we love and admire the most. . . . Emulation, the anxious desire that we ourselves should excel, is originally founded in our admiration of the excellence of others. Neither can we be satisfied with being merely admired for what other people are admired. We must at least believe ourselves to be admirable for what they are admirable. . . . [If others praise us, their] praise necessarily strengthens our own sense of our own praise-worthiness. In this case, so far is the love of praise-worthiness from being derived altogether from that of praise; that the love of praise seems, at least in a great measure, to be derived from that of praise-worthiness.<sup>14</sup>

In order to accommodate this observation, Smith postulates a primitive desire to be worthy of praise.

Nature, when she formed man for society, endowed him with an original desire to please, and an original aversion to offend his brethren. She taught him to feel pleasure in their favourable, and pain in their unfavourable regard. . . .

But this desire of the approbation, and this aversion to the disapprobation of his brethren, would not alone have rendered him fit for that society for which he was made. Nature, accordingly, has endowed him, not only with a desire of being approved of, but with a desire of being what ought to be approved of; or of being what he himself approves of in other men.<sup>15</sup>

This passage is difficult to interpret, since a desire “of being what ought to be approved of” is not quite the same as a desire “of being what [one] approves of in other men.” And, simply postulating a desire to be “what ought to be approved of” would not eliminate the difficulty. According to Smith’s impartial spectator theory, a desire to be “what ought to be approved of” is precisely a desire to be such that impartial spectators *would* approve of oneself. But, as Smith emphasizes at length, that desire is not yet of the right sort.

A desire to be “what he himself approves of in other men” is more to the point, if it means a desire to be “what he himself would approve of in others from an impartial perspective.” But Smith does not really postulate any such desire as a basic unexplained fact about people. Instead, he offers a more complex account of moral motivation that anticipates certain elements of Freud’s theory of the super-ego. Smith holds that, at first, a child is motivated to restrain its reactions so as to have the sympathy of parents and other spectators. As an aid in doing this, the child tries to view itself as seen by others. Eventually, it acquires a habit of doing this—a habit of pretending to be an impartial spectator of its own actions and reactions to see which actions it approves and which it disapproves. The child then tries to restrain its reactions so as to

be approved by this pretend spectator—the little person within who now serves as a kind of inner conscience.

When I endeavour to examine my own conduct, when I endeavour to pass sentence upon it, and either to approve or condemn it, it is evident that, in all such cases, I divide myself, as it were, into two persons; and that I, the examiner and judge, represent a different character from that other I, the person whose conduct is examined into and judged of. The first is the spectator, whose sentiments with regard to my own conduct I endeavour to enter into, by placing myself in his situation, and by considering how it would appear to me, when seen from that particular point of view. The second is the agent, the person whom I properly call myself, and of whose conduct, under the character of a spectator, I was endeavouring to form some opinion. The first is the judge; the second the person judged of. But that the judge should, in every respect, be the same with the person judged of, is as impossible as that the cause should, in every respect, be the same with the effect.<sup>16</sup>

This might seem to be even worse than purely outer directed motivation. It may seem that Smith is saying that the agent is motivated to obtain the sympathy and approval of a *pretend* person.<sup>17</sup>

But this is a mistake. Despite the division into what Smith calls “two persons”, the “examiner and judge” is not really any different from “the person I call myself.” The “examiner and judge” is the agent himself or herself, viewing things from a certain perspective. When the agent pretends to be an impartial spectator, he or she ends up approving and disapproving from an impartial point of view. Viewing things in that way can then influence the agent’s motives and feelings, since it is the agent who is doing the viewing and the approving and disapproving and the approval or disapproval is of the agent himself or herself. Consequently, the agent’s actual motives will become more moral, because they are in part the result of the agent’s looking at things from a moral point of view. To the extent that the agent views things impartially, the agent will genuinely not want to cheat and injure others. What starts as a strategy for knowing what to do to obtain sympathy ends up giving the agent a genuinely moral outlook that can motivate moral behavior.

The truly moral agent does not aim at getting the sympathy of impartial observers. Instead, the moral agent is motivated to act morally by virtue of motives acquired by viewing things from an impartial standpoint. The moral agent cares about the things that an impartial spectator cares about because the agent has in part *become* an impartial spectator.

### *Concluding Summary*

Hutcheson, Hume, and Smith all put forward versions of the impartial spectator theory. All three agree that right acts are right because they would be favored by impartial spectators who favor these acts for other reasons than that the acts are right. But they disagree about what explains spectators' reactions and what accounts for moral motivation. They also disagree as to whether the same thing explains agents' motivation and spectators' approval.

Hutcheson and Hume agree that spectators' reactions derive from benevolence. Hutcheson believes that benevolence is also the source of moral motivation. Hume thinks this is too weak a motivation; he thinks moral motivation is also based in self-interest. Where Hutcheson simply treats benevolence as a God given motive, Hume offers an explanation of benevolence in terms of what he calls sympathy.

Smith disagrees with Hume as to what sympathy is. Smith also disagrees with both Hume and Hutcheson about the relevance of benevolence. Smith argues that neither agents nor spectators are much influenced by benevolence, although the desire for sympathy is important. Smith agrees with Hutcheson and disagrees with Hume in holding that moral motivation is of a piece with what it is that leads spectators to approve or disapprove of various actions. What is important, according to Smith, is the taking of an impartial view.

Smith works this theory out with a mass of detail which I cannot try to summarize. I believe that the book in which he works this out, his *Theory of the Moral Sentiments*, is one of the great works of moral philosophy.

I conclude that the Impartial Spectator Theory has an answer to the objection that it offers an overly aesthetic conception of morality, that it overemphasizes the point of view of the spectator over the point of view of the agent, and that it cannot account for the motivation to do what is right.

This is not to say that I am convinced that this is the best theory. One alternative is to try to develop what might be called an ideal agent or ideal practical reasoner theory, as opposed to an ideal observer or impartial spectator theory. This alternative would place primary emphasis on moral reasons for acting, on the viewpoint of the agent. In one version of this theory, the agent has moral principles that the agent intends to follow. The agent's moral reasons typically derive from principles the agent accepts as a member of a group.<sup>18</sup>

I am not sure how to decide between the impartial spectator theory and the ideal reasoner theory.

So let me conclude with a couple of further comments on the history of moral philosophy in the 18th Century.

### *Postscript 1*

There is an interesting irony in the way in which Hume's use of the term "sympathy" leads Smith to his own very different theory, a theory that in my view is much better than Hume's at accounting for moral phenomenology. Smith's criticism of Hume's use of the term "sympathy" is not a serious one. It is of no importance whatsoever whether the meaning that Hume gives to the term "sympathy" is the ordinary one and it in no way damages Hume's view for him to acknowledge that his use is nonstandard. Hume can use the term however he wants. If he wants to use it in a special sense in order to develop his own view, there is nothing wrong with that. If someone really objects, the term can be replaced with another. Nothing in Hume's view depends on his having captured the ordinary meaning of the term "sympathy." The irony is that taking Hume's term seriously leads Smith to a more accurate account of morality. A purely verbal point yields a powerful substantive theory.

### *Postscript 2*

Finally, it is perplexing that Adam Smith's ethics should be so relatively unread as compared with Hume's ethics when there is so much of value in Smith. What I have talked about here only scratches the surface. Why should Smith's ethics be so neglected? Is it that Hume also had a metaphysics and an epistemology and that Smith did not? Or is it that Smith was a more important economist than Hume? And why should that matter? I do not know.<sup>19</sup>

## NOTES

1. Francis Hutcheson, *An Inquiry concerning Moral Good and Evil* (1725) and *Illustrations upon the Moral Sense* (1728). David Hume, *A Treatise of Human Nature* (London 1739-40) and *Enquiries concerning Human Understanding and the Principles of Morals* (London 1748-1751). Adam Smith, *Theory of the Moral Sentiments*, first edition 1759, sixth edition, 1790. References below are to the Glasgow edition, Oxford 1976, reprinted in paper covers by Liberty Press, Indianapolis, Indiana, 1982.

2. Roderick Firth, "Ethical Absolutism and the Ideal Observer," *Philosophy and Phenomenological Research* 12 (1952). Richard B. Brandt, *Ethical Theory* (Englewood Cliffs, New Jersey; Prentice-Hall: 1959). F. C. Sharp, *Good Will and Ill Will* (Chicago, Chicago University Press: 1950).

3. David Wiggins, John McDowell, and Thomas L. Carson have argued in favor of a view of this sort. David Wiggins, "Truth, Invention, and the Meaning of Life" (British Academy Lecture: 1976); "A Sensible Subjectivism," ms. John McDowell, "Values and Secondary Qualities," in Ted Honderich, *Morality and Objectivity: A Tribute to J. L. Mackie*

(London, Routledge and Kegan Paul: 1985), pp. 110-129. Thomas L. Carson, *The Status of Morality* (Dordrecht, Reidel: 1984). Colin McGinn and Simon Blackburn have argued against it. Simon Blackburn, *Spreading the Word* (Oxford, Oxford University Press: 1984), chapter 6; "Errors and the Phenomenology of Value," in Honderich (ed.) *op. cit.*, pp. 1-22. Colin McGinn, *The Subjective View* (Oxford, Oxford University Press: 1983).

4. Stuart Hampshire, "Fallacies in Moral Philosophy," in *Freedom of Mind and Other Essays* (Oxford, Oxford University Press: 1972).

5. *The Nature of Morality: An Introduction to Ethics* (New York, Oxford University Press: 1977).

6. Firth, *op. cit.*

7. Robert Shope notes that problems of this sort arise almost always when philosophers appeal to hypothetical cases in this way. Shope, "The Conditional Fallacy," *Journal of Philosophy* 75 (1978) 397-413.

8. Francis Hutcheson, *An Inquiry concerning Moral Good and Evil*, 2nd edition (London: 1726), III, viii (paragraph 121 in *British Moralists*, edited by L. A. Selby-Bigge (Oxford, Oxford University Press: 1897—paper edition Indianapolis, Indiana; Bobbs-Merrill: 1964)).

9. Thomas Hobbes, *Leviathan, or the Matter, Form and Power of a Commonwealth* (London: 1651).

10. Bernard de Mandeville, *The Fable of the Bees, or Private vices publick benefits* (1714).

11. Immanuel Kant, *Foundations of the Metaphysics of Morals* (1785).

12. Smith, *op. cit.* I,i,1,10.

13. *Ibid.* I,i,3,2.

14. *Ibid.* III.2.2-3. Similar ideas have been recently expressed by John McDowell, *op. cit.*, pp. 117-120 and by David Wiggins, "A Sensible Subjectivism," ms p. 4.

15. *Ibid.* III.2.6-7.

16. *Ibid.* III,1,6.

17. I mistakenly offer this as a criticism of basing morality on the Freudian super-ego in Harman, *op. cit.*, pp. 61-62.

18. Harman, *The Nature of Morality*, Chapter 9.

19. I am indebted to David Levy for getting me interested in this project, to Ralph Lindgren's *Social Philosophy of Adam Smith* (The Hague, Martinus Nijhoff: 1973), which persuaded me that my initial thoughts about Smith were superficial, and to Michael Smith for helpful comments on a prior draft. (This is not to say that any of them agree with what I have said.) Earlier versions of this paper were presented at a Liberty Fund conference on Adam Smith in Washington, as the Selfridge Lecture at Lehigh University in 1984, as the Matchette Lecture at Trinity College (San Antonio, Texas) and at the University of Miami.

The following lectures have been published in individual pamphlet form and may be obtained from the Department at a price of one dollar and fifty cents each.

- \*1961. "The Idea of Man—An Outline of Philosophical Anthropology."  
By José Ferrater Mora, Professor of Philosophy, Bryn Mawr College.
- 1962. "Changes in Events and Changes in Things."  
By A. N. Prior, Professor of Philosophy, University of Manchester.
- \*†1963. "Moral Philosophy and the Analysis of Language."  
By Richard B. Brandt, Professor of Philosophy, Swarthmore College.
- \*†1964. "Human Freedom and the Self."  
By Roderick M. Chisholm, Professor of Philosophy, Brown University.
- †1965. "Freedom of Mind."  
By Stuart Hampshire, Professor of Philosophy, Princeton University.
- \*†1966. "Some Beliefs about Justice."  
By William K. Frankena, Professor of Philosophy, University of Michigan.
- †1967. "Form and Content in Ethical Theory."  
By Wilfrid Sellars, Professor of Philosophy, University of Pittsburgh.
- †1968. "The Systematic Unity of Value."  
By J. N. Findlay, Clark Professor of Philosophy, Yale University.
- 1969. "Buber and Buberism—A Critical Evaluation."  
By Paul Edwards, Professor of Philosophy, Brooklyn College of the City University of New York.
- 1971. "What Actually Happened."  
By P. H. Nowell-Smith, Professor of Philosophy, York University.
- †1972. "Moral Rationality."  
By Alan Gewirth, Professor of Philosophy, University of Chicago.
- †1973. "Reflections on Evil."  
By Albert Hofstadter, Professor of Philosophy, University of California, Santa Cruz.
- ††1974. "What is Dialectical?"  
By Paul Ricoeur, Professor of Philosophy, University of Paris and University of Chicago.
- ††1975. "Some Confusions About Subjectivity."  
By R. M. Hare, White's Professor of Moral Philosophy at Oxford University and Fellow of Corpus Christi College.
- 1976. "Self-Defense and Rights."  
By Judith Jarvis Thomson, Professor of Philosophy, Massachusetts Institute of Technology.
- 1977. "What is Humanism?"  
By Georg Henrik von Wright, Research Professor of Philosophy, The Academy of Finland.
- 1978. "Moral Relativism."  
By Philippa Foot, Senior Research Fellow, Somerville College, Oxford; and Professor of Philosophy, University of California, Los Angeles.
- 1979. "The Idea of the Obscene."  
By Joel Feinberg, Professor of Philosophy, University of Arizona.
- 1980. "Goods Beyond Price and Other Apparent Anachronisms."  
By Warner Wick, Professor of Philosophy, University of Chicago.
- 1981. "Morality, Property and Slavery."  
By Alan Donagan, Professor of Philosophy, University of Chicago.
- 1982. "Expressing Evaluations."  
By Donald Davidson, Professor of Philosophy, University of California, Berkeley.

(continued, back cover)

\* Pamphlet out of print.

† Reprinted in *Freedom and Morality*.

†† Printed only in *Freedom and Morality*.

1983. "How Not to Solve Ethical Problems."  
By Hilary Putnam, Professor of Philosophy, Harvard University.
1984. "Is Patriotism a Virtue?"  
By Alasdair MacIntyre, W. Alton Jones Professor of Philosophy,  
Vanderbilt University.
1985. "How Free Does the Will Need to Be?"  
By Bernard Williams, Provost of King's College Cambridge.

**A Volume of Lindley Lectures  
Delivered at the University of Kansas**

# Freedom and Morality

**Richard B. Brandt**

Moral Philosophy and the  
Analysis of Language

**Roderick M. Chisholm**

Human Freedom and the Self

**Stuart Hampshire**

Freedom of Mind

**William K. Frankena**

Some Beliefs about Justice

**Wilfrid Sellars**

Form and Content in Ethical  
Theory

**J. N. Findlay**

The Systematic Unity of Value

**Alan Gewirth**

Moral Rationality

**Albert Hofstadter**

Reflections on Evil

**Paul Ricoeur**

What Is Dialectical?

**R. M. Hare**

Some Confusions about  
Subjectivity

Edited with an Introduction by

**John Brice**

The volume can be purchased for  
**\$6.00** from the Library Sales  
Office, University of Kansas  
Libraries, Lawrence, Kansas  
(U.S.A.) 66045. Please include a  
50¢ handling fee, 75¢ outside the  
United States.