XIV*—MORAL PHILOSOPHY MEETS SOCIAL PSYCHOLOGY

Virtue Ethics and the Fundamental Attribution Error

by Gilbert Harman

ABSTRACT Ordinary moral thought often commits what social psychologists call 'the fundamental attribution error'. This is the error of ignoring situational factors and overconfidently assuming that distinctive behaviour or patterns of behaviour are due to an agent's distinctive character traits. In fact, there is no evidence that people have character traits (virtues, vices, etc.) in the relevant sense. Since attribution of character traits leads to much evil, we should try to educate ourselves and others to stop doing it.

I

Folk physics and folk morality. Ordinary untrained physical intuitions are often in error. For example, ordinary people expect that something dropped from a moving vehicle or airplane will fall straight down to the point on earth directly underneath the place from which it was released. In fact, the dropped object will fall in a parabolic arc in the direction of the movement of the vehicle or airplane from which it was dropped. This means, among other things, that bombardiers need to be trained to go against their own physical intuitions. There are many similar examples (McCloskey, 1983; Holland, Holyoak, Nisbett, & Thagard, 1986).

Considering the inadequacies of ordinary physical intuitions, it is natural to wonder whether ordinary moral intuitions might be similarly inadequate. And, while many moral philosophers seem to put great confidence at least in their own moral intuitions, others argue for revisions. Consequentialism may be put forward not as an attempt to capture intuitive folk morality but rather as a critique of ordinary intuitions (Kagan, 1989). Similarly, moral relativism might be defended as the truth about

*Meeting of the Aristotelian Society, held in Senate House, University of London, on Monday, 21st June, 1999 at 8.15 p.m.

morality, whether or not moral relativism accords with everyone's intuitions (Harman, 1996).

On this occasion I discuss a different kind of rejection of folk morality, one that derives from contemporary social psychology. It seems that ordinary attributions of character traits to people are often deeply misguided and it may even be the case that there is no such thing as character, no ordinary character traits of the sort people think there are, none of the usual moral virtues and vices.

In attempting to characterise and explain the movements of a body, folk physics places too much emphasis on assumed internal characteristics of the body, ignoring external forces. Similarly, in trying to characterise and explain a distinctive action, ordinary thinking tends to hypothesise a corresponding distinctive characteristic of the agent and tends to overlook the relevant details of the agent's perceived situation.[1] Because of this tendency, folk social psychology and more specifically folk morality are subject to what Ross (1977) calls 'the fundamental attribution error'.

Empirical studies designed to test whether people behave differently in ways that might reflect their having different character traits have failed to find relevant differences. It is true that studies of this sort are very difficult to carry out and there have been very few such studies. Nevertheless, the existing studies have had negative results. Since it is possible to explain our ordinary belief in character traits as deriving from certain illusions, we must conclude that there is no empirical basis for the existence of character traits.

II

Character. Character traits must be distinguished from psychological disorders like schizophrenia, mania, and depression, and from innate aspects of temperament such as shyness or being basically a happy or sad person. Character traits include virtues and vices like courage, cowardice, honesty, dishonesty, benevolence, malevolence, friendliness, unfriendliness, as well as certain other traits like friendliness or talkativeness.

1. 'The relation between lay personology and a more correct theory of personality is analogous to the relation between lay and scientific physics', Ross and Nisbett, 1991: 161, citing earlier work including Lewin (1935).

Aristotle (1985) describes the ordinary conception of such character traits. They are relatively long-term stable disposition to act in distinctive ways. An honest person is disposed to act honestly. A kind person is disposed to act kindly. The relevant dispositions must involve habits and not just skills, involving habits of desiring. To be sure, as we normally conceive of certain character traits or virtues, they may involve certain strengths or skills, as in courage or strength of will (Brandt, 1988). But they involve more than simply having relevant skills or know-how. A person with the relevant character trait has a long term stable disposition to use the relevant skills in the relevant way. Similarly, the virtue of benevolence may involve practical knowledge concerning how to benefit people; but mere possession of that knowledge with no disposition to use it to benefit people would be insufficient for possession of a benevolent character.

In ordinary conceptions of character traits and virtues, people differ in their possession of such traits and virtues. A particular character trait fits into one or more ranges of ways of behaving. In some cases, the relevant virtue can be seen as a mean between extremes (Aristotle, 1985). Courage is a mean between rashness and timidity, for example. Proper benevolence is a mean between stinginess and profligacy. Where some people have a given virtue, others have one or another corresponding vice. Different ways in which people behave on different occasions are sometimes due to their having such different character traits. Finding a wallet on the sidewalk, an honest person tries to locate the owner, whereas a dishonest person pockets the contents and throws the rest of the wallet away. How a stranger reacts to you depends whether the stranger is basically friendly or unfriendly.

We ordinarily suppose that a person's character traits help to explain at least some things that the person does. The honest person tries to return the wallet *because* he or she is honest. The person who pockets the contents of the wallet and throws the rest of the wallet away does so *because* he or she is dishonest.

The fact that two people regularly behave in different ways does not establish that they have different character traits. The differences may be due to their different situations rather than differences in their characters. To have different character traits, they must be disposed to act differently in the same circumstances (as they perceive those circumstances).

Furthermore, character traits are *broad based* dispositions that help to explain what they are dispositions to do. Narrow dispositions do not count. If fifteen year old Herbert is disposed to refuse to ride any roller coaster, but is not cowardly or fearful in other ways, his particular disposition is not an instance of cowardice or fear and indeed may fail to be an instance of any character trait at all. If Herbert also acquires a disposition to refrain from speaking up in history class (but not in other subjects) and the explanation of this latter reluctance is quite different from the explanation of his avoidance of roller coaster rides, then these two dispositions are not special cases of a single character trait. Nor can cowardice or fearfulness be constructed out of a collection of quite separable dispositions of this sort, if there is no common explanation of the resulting behaviours.

III

Virtue Ethics. Some theorists suppose that proper moral development requires moral instruction in virtue.[2] In this view, moral instruction involves teaching relevant habits of action, perhaps habits of desire, in some cases also relevant skills. If a learner's dispositions fall more toward one of the extremes in one or another relevant range of behaviour, moral educators should encourage the learner to aim more towards the opposite extreme until the right balance is achieved. It is occasionally remarked that one thing wrong with contemporary American society is that too little attention is being paid to this sort of character development (e.g., Bennett, 1993).

Some philosophers argue, further, that morality or perhaps the ordinary conception of morality is best analyzed by beginning with a conception of virtue and character and then explaining other aspects of morality in terms of that (Taylor, 1991; Hursthouse, 1996). In this view, we determine what a person *ought morally* to do in a particular situation by considering what a person of good character would do in that situation. An act is *morally right* to the extent that it is the result of the agent's good character and *morally wrong* to the extent that it is the result of

2. An alternative view (Harman, forthcoming) is that children no more require moral instruction in order to acquire morality than they require instruction in their first language in order to acquire that language.

the agent's bad character. Perhaps we can also say that a situation or state of affairs is morally good to the extent that it would be favoured by a good person.

Some versions of virtue ethics connect virtues with human flourishing. In one version, a virtue is a character trait that contributes to the flourishing of the agent. In another version, the virtues are character traits that contribute to the flourishing of people in general. In either version, it is not easy to provide a noncircular account of human flourishing that leaves the resulting view sounding plausible (Harman, 1983).

The details of how virtue ethics might be developed are interesting, but I do not want to get into them on this occasion. For present purposes, the main point is that this sort of virtue ethics presupposes that there are character traits of the relevant sort, that people differ in what character traits they have, and these traits help to explain differences in the way people behave.[3]

IV

Social Psychology. Philosophers have begun to notice that recent social psychology challenges ordinary and philosophical views about character traits. Flanagan (1991) discusses the challenge at length, arguing that it is not as radical as it may seem. Railton (1997) thinks the challenge is more serious, as does Doris (forthcoming) in an important book length study.

Let me begin my own account by emphasising that the empirical results of contemporary social psychology can seem extremely counter-intuitive on first acquaintance. Students of mine who read parts of Nisbett and Ross' useful textbook (Nisbett & Ross, 1991) report that their parents express dismay at the 'nonsense' they are being taught at Princeton.

Flanagan (1991), who is a philosophical pioneer in discussing the relevant social-psychological literature, does not seem to me fully to appreciate its radical import. He mentions what he calls the 'extreme view', according to which 'Good behaviour is not the result of good character. It is the result of a certain kind of dominating environment. Take away the powerful external

3. But see VII.1.1 below, where I mention two versions of 'virtue ethics' that do not treat virtues as traits of character.

props, and what seems to be a consistently good character will evaporate into thin air'. He continues, 'Almost no one holds such an extreme view'. However, contrary to this remark of Flanagan's, the 'extreme view' is in fact widespread among social psychologists.

Nisbett and Ross (1991) report that '[t]he experience of serious graduate students, who, over the course of four or five years, are immersed in the problems and the orientation of the field [of social psychology],... is an intellectually wrenching one. Their most basic assumptions about the nature and the causes of human behaviour...are challenged' (1).

At one point, Nisbett and Ross 'seriously entertained the hypothesis that most of [the] seeming order [in ordinary human behaviour] was a kind of cognitive illusion. We believed that human beings are adept at seeing things as they believe them to be, at explaining away contradictions and, in particular, at perceiving people as more consistent than they really are'. Nisbett and Ross now think that there are at least regularities in human behaviour and that lay personality may work in the sense of enabling people to manage in ordinary life, just as lay physics works for many ordinary situations. 'That is, people often make correct predictions on the basis of erroneous beliefs and defective prediction strategies' (7–8).

> [I]n everyday experience the characteristics of actors and those of the situations they face are typically confounded—in ways that contribute to precisely the consistency that we perceive and count on in our social dealings. People often choose the situations to which they are exposed; and people often are chosen for situations on the basis of their manifest or presumed abilities and dispositions. Thus, clerics and criminals rarely face an identical or equivalent set of situational challenges. Rather they place themselves, and are placed by others, in situations that differ precisely in ways that induce clergy to look, act, feel, and think rather consistently like clergy and that induce criminals to look, act, feel, and think like criminals (19).

In addition, 'individuals may behave in consistent ways that distinguish them from their peers not because of their enduring predispositions to be friendly, dependent, aggressive, or the like, but rather because they are pursuing consistent goals using consistent strategies, in the light of consistent ways of interpreting

their social world' (20). And 'people sometimes feel *obliged*, even committed to act consistently. This may be because of their social roles, because of the real-world incentives' etc. (19).

V

Two Experiments. Social psychologists have shown many different ways in which ordinary observers wrongly infer that actions are due to distinctive character traits of an agent rather than relevant aspects of the situation. Here I briefly review two well known experiments, one by Millgram and one by Darley and Batson.

V.1 Obedience to Authority

Millgram (1963) describes an experiment in which a subject was given the task of administering an increasingly intense electric shock to a second person, the 'learner', whenever the learner gave the wrong answer. (Subjects were also told to treat a failure to answer as a wrong answer.) The shocks started at 15 volts and increased in 15 volt intervals to the highest level of 450 volts. The device used had labels at various points indicating 'Slight Shock', 'Moderate Shock', 'Strong Shock', 'Very Strong Shock', 'Intense Shock', 'Extreme Intensity Shock', 'Danger: Severe Shock', and 'XXX'. At the 300 volt level the learner pounded loudly on the wall of the room but did not answer the question. This is repeated at the 315 volt level. At higher levels there was no further response from the learner.

Whenever the subject asked the experimenter for advice or the subject said he did not want to continue, the experimenter had a list of four things to say, which would be said only if needed and only in sequence: (1) 'Please continue' or 'Please go on'. (2) 'The experiment requires that you continue'. (3) 'It is absolutely essential that you continue', and (4) 'You have no other choice, you *must* go on'. If the subject persisted in asking to stop after being told these four things, he or she would then be excused.

The experiment was designed to test how far subjects would go in administering shock under these conditions. The experimenters had expected that few subjects would go beyond the designation 'Very Strong Shock' (150 volts). But in fact, of the 40 subjects in one (typical) early experiment, *all* went past that

point. Five stopped at the 300 volt level right before the label 'Extremely Intense Shock' and the point at which the learner pounded on the wall. Four more stopped at the next stage, 315 volts, when the learner pounded the wall again. Two stopped at 330 volts, when the learner made no response at all. One stopped at 345 volts and another at 360 volts. The 26 remaining subjects, 65 per cent of the total, continued on to 450 volts. In other words, most of the 40 subjects went all the way to give the maximum shock.

To repeat an important point, the experimenters (and others whom they questioned both before and after) did not at all expect this sort of result. They expected almost everyone to stop well before 300 volts, by 150 volts. In addition, people who have had the experiment described to them in detail, tend to be quite confident that, if they had participated in the original experiment, they would have stopped administering shocks at or before that relatively early point (150 volts), much earlier than anyone did in the actual experiment.

Now consider any one of the subjects who went all the way to 450 volts, past the label 'Danger: Severe Shock' and well past the point at which the learner had stopped responding in any way. It is hard not to think there is something terribly wrong with the subject. It is extremely tempting to attribute the subject's performance to a character defect in the subject rather than to details of the situation.

But can we really attribute a 2 to 1 majority response to a character defect? And what about the fact that *all* subjects were willing to go at least to the 300 volt level? Does *everyone* have this character defect? Is that really the right way to explain Millgram's results?

A different kind of explanation (Ross & Nisbett, 1991: 56–8) invokes relevant features of the situation. First, there is 'the stepwise character of the shift from relatively unobjectionable behaviour to complicity in a pointless, cruel, and dangerous ordeal', making it difficult to find a rationale to stop at one point rather than another. Second, 'the difficulty in moving from the intention to discontinue to the actual termination of their participation', given the experimenter's refusal to accept a simple announcement that the subject is quitting—'The experiment requires that you continue'. Third, as the experiment went on, 'the events that

unfolded did not "make sense" or "add up" .... The subjects' task was that of administering severe electric shocks to a learner who was no longer attempting to learn anything .... [T]here was simply no way for [subjects] to arrive at a stable "definition of the situation".'

The fundamental attribution error in this case consists in 'how readily the observer makes erroneous inferences about the actor's destructive obedience (or foolish conformity) by taking the behaviour at face value and presuming that extreme personal dispositions are at fault.'

V.2 Good Samaritans

The second experiment that I will mention derives from the parable of the Good Samaritan, which goes like this.

> 'And who is my neighbour?' Jesus replied. 'A man was going down from Jerusalem to Jericho, and he fell among robbers, who stripped him and beat him, and departed, leaving him half dead. Now by chance a priest was going down the road; and when he saw him he passed by on the other side. So likewise a Levite, when he came to the place and saw him, passed by on the other side. [Levites were important participants in temple ceremonies.] But a Samaritan [a religious outcast], as he journeyed, came to where he was; and when he saw him, he had compassion and went to him and bound his wounds, pouring on oil and wine; then he set him on his own beast and brought him to an inn, and took care of him. And the next day he took out two dennarii and gave them to the innkeeper, saying, 'Take care of him; and whatever more you spend, I will repay you when I come back.' Which of these three, do you think, proved neighbour to him who fell among the robbers? He said, 'The one who showed mercy on him.' And Jesus said to him, 'Go and do likewise.' (Luke 10: 29-37, Revised Standard Version.)

Darley and Batson (1973) observe that people can envision various differences between the priest and Levite on the one hand and the Samaritan on the other hand. The priest and Levite might have well have had their minds on religious matters, whereas the Samaritan probably did not. The priest and Levite were probably hurrying along to various appointments, whereas the Samaritan was probably less in a hurry. The parable also

suggests that there is a difference in type of religiosity or morality. The priest and Levite in Jesus's act virtuously in order to please God, where the Samaritan responds more directly to the needs of another person.

The standard interpretation of the parable focuses on the third of these variables, the type of religious or moral character of the agent.

Darley and Batson designed an experiment aimed at uncovering which of these differences might be most relevant to explaining the differences in behaviour. Subjects in this experiment were students at Princeton Theological Seminary. As each subject arrived, he was informed that he was to give a talk that would be recorded in another building. Along the way to the place for the talk, the subject encountered a 'victim' slumped in a doorway. The question was under what conditions would a subject stop to help the victim.

Half of the subjects were assigned to talk on the Good Samaritan Parable; the others were assigned a different topic. Some of the subjects were told they were late and should hurry; some were told they had just enough time to get to the recording room; and some were told they would arrive early. Judging by their responses to a questionnaire, they had different religious and moral orientations.

The only one of these variables that made a difference was how much of a hurry the subjects were in. 63 per cent of subjects that were in no hurry stopped to help, 45 per cent of those in a moderate hurry stopped, and 10 per cent of those that were in a great hurry stopped. It made no difference whether the students were assigned to talk on the Good Samaritan Parable, nor did it matter what their religious outlook was.

Standard interpretations of the Good Samaritan Parable commit the fundamental attribution error of overlooking the situational factors, in this case overlooking how much of a hurry the various agents might be in.

VI

Direct Empirical Challenges to Character Traits. But don't we know from ordinary experience that people differ in character traits? Here it is useful to consider related issues.

Psychoanalysts acquire a considerable experience in treating patients and can cite many instances in which psychoanalytic treatment is successful. However, empirical studies of psychoanalytic treatment as compared with no treatment have found no objective benefit. (Dawes, 1994).

Some diagnosticians have used Rorschach inkblot tests to make psychological diagnoses. It seemed to those using these tests that they had abundant evidence that certain characteristics of the test results were diagnostic of certain disorders. Empirical studies showed there was no correlation between those characteristics and the test results. (Nisbett & Ross, 1980: 93–7).

Many employers are convinced that useful information can be gained from interviewing potential employees. However, for the most part, interviews simply add noise to the decision process. Empirical studies indicate that decisions made on information available apart from an interview are more reliable than decisions made when an interview is added. (Ross & Nisbett, 1991: 136–8).

Discovery of such errors in reasoning has encouraged research into why people are subject to such errors (Tversky & Kahneman, 1974, Nisbett & Ross, 1980). One suggested reason is confirmation bias. Given a hypothesis, one tends to look for confirming evidence. Finding such evidence, one takes it to support the hypothesis. Evidence against the hypothesis tends to be ignored or downplayed.

Ross and Nisbett suggest that the initial source of the fundamental attribution error may have to do with Gestalt considerations of figure and ground. Where we distinguish figure from ground, we pay more attention to figure and less to ground and we try to explain what happens in terms of features of the figure rather than in terms of features of the ground. Typically the actor is figure and the situation is ground, so we seek an explanation of the action in features of the actor in the foreground rather than in features of the background situation. The suggested explanation is then subject to confirmation bias. Additional support comes from the fact that other people give explanations in terms of dispositional features of agents rather than in terms of aspects of their situations.

When investigators have looked for objective evidence that people differ in character traits, the results have been much as

with psychoanalysis, Rorschach tests, and interviews. People take themselves to have lots of evidence that agents differ in character traits. Yet empirical studies have not found any objective basis for this confidence. Summarising a number of studies, Ross and Nisbett (1991: 95) report that the 'average correlation between different behavioural measures designed to tap the same personality trait (for examples, impulsivity, honesty, dependency, or the like) was typically in the range between 0.10 and 0.20, and often was even lower'. These are very low correlations, below the level which people can detect. Using such correlations to make predictions yields hardly any improvement over guessing. Even if predictions are limited to people one takes to be quite high on a particular trait, the correlations are still very low.

Ross and Nisbett observe that people have some appreciation of the role of situation in the way they understand such stories as *The Prince and the Pauper* or the movie *Trading Places*. But for the most part, people are quick to infer from specific actions to character traits.

It is true that there are better correlations for very specific situations. 'Hartshorne and May (1928) found that the tendency to copy from an answer key on a general information test on one occasion was correlated 0.79 with copying from an answer key on a similar test six months later. Newcomb (1929) found that talkativeness at lunch was a highly stable attribute; it just was not very highly correlated with talkativeness on other occasions....' (Ross & Nisbett, 1991: 101).

Surprisingly, Flanagan (1991) argues that this shows there really are character traits, 'albeit not traits of unrestricted globality or totally context-independent ones'. I guess he means such character traits as 'being disposed to copy from an answer key on a certain sort of test' and 'being talkative at lunch'. But, first, no reason has been given for thinking that these specific narrow regularities in behaviour reflect dispositions or habits rather than, for example, skills or strategies that have worked in the past. Second, and more importantly for our purposes, ordinary thinking about personality and character attributes is concerned with more global traits like honesty and talkativeness.

Flanagan concludes: 'Yes, there are character traits. The language of character traits picks out psychologically real phenomena.' But I do not see that he has cited any empirical evidence for this claim.

Flanagan also seems to think that it inconsistent to argue against character traits by appeal to the fundamental attribution error. He says, 'It is telling against the situationalist who is also an eliminativist that he will have extreme difficulty (indeed he courts inconsistency) in positing attributional biases of any sort if by these he means to refer to what he must be taken to want to refer to, namely, *dispositions* to think in certain ways' (305). But this is true only if a 'situationalist' is someone who denies that there are any dispositions at all, or who (perhaps like Skinner, 1974) denies that it is useful to explain anything in terms of dispositions. The issue we have been concerned with is whether people differ in certain particular dispositions—character traits. To deny that people differ significantly in character traits is not to deny that they have any dispositions at all. People might well all share certain dispositions, such as a disposition to make the fundamental attribution error. Secondly, they might differ in various dispositions that do not constitute character traits, such as personality disorders and other mental illnesses. (So, for example, to deny that there are character traits is not to accept the view in Laing, 1960, that schizophrenia is simply a rational response to a difficult family situation.)

VII

Benefits of Appreciating the Fundamental Attribution Error. There are various benefits to a proper appreciation of ways in which ordinary moral thinking rest on the fundamental attribution error.

VII.1 Philosophy

VII.1.1 Virtue Ethics

Character based virtue ethics may offer a reasonable account of ordinary moral views. But to that extent, these ordinary views rest on error.

It is worth mentioning that there are variants of virtue ethics that do not require character traits in the ordinary sense. For example, Thomson (1996) tries to explicate moral thinking by appeal to judgements about whether particular actions are just or courageous or whatever. To the extent that such judgements

are concerned entirely with the action and not with any presumed underlying trait of character, Thomson's enterprise is unaffected by my discussion.

Maria Merritt (forthcoming) has been developing a version of virtue ethics that emphasises the role of the situation in maintaining relevant regularities in behaviour.

VII.1.2 Better Understanding of Moral Luck

Adam Smith (1759) wrote about the influence of fortune on our moral judgements, giving nice examples. Someone carelessly throws a brick over a wall. His companion may complain about this even if no harm is done. But if the brick does hit someone, much greater condemnation ensues. Nagel (1979) gives a similar example of a driver who takes his eyes off the road for a second. That's bad, but suppose in that second a child darts into the street and is hit. Then much worse condemnation seems appropriate.

Smith and Nagel note that from a certain point of view, our moral judgement of the act should be based entirely on the motives of the agent and the agent's epistemic situation, so that from that point of view there should be no difference between two cases that are the same in those respects in one of which someone is hit by the brick (or car) and in the other of which no one is hit. Yet, it is clear that we will judge the cases differently.

Perhaps these are simply further instances of the fundamental attribution error. This bad thing has happened and we attribute it to the bad character of the agent in the foreground.

VII.2 Real Life

VII.2.1 Moral Education

If there is no such thing as character, then there is no such thing as character building.

VII.2.2 Tolerance

When things go wrong, we typically blame the agent, attributing the bad results to the agent's bad character. Even when things do not go bad, we are quick to interpret actions as expressive of character traits, often hostile traits. For example, a person with

poor vision may fail to recognise an acquaintance, who then attributes this to coldness in that person.

A greater understanding of the agent's situation and how it contributed to the action can lead to a greater tolerance and understanding of others.

VII.2.3 Better Understanding of Ethnic Hatred

Recent terrible events in the former Yugoslavia are often attributed to historical 'ethnic hatreds'. Yet it is possible to explain these events in rational terms (Hardin, 1995). Suppose there are limited resources and a successful coalition will benefit its members more than those excluded from the coalition. Such a coalition is possible only if insiders can be distinguished from excluded outsiders and only if it is possible to keep members from defecting to other groups. Coalitions formed around ethnic or religious lines might succeed. The threat that one such coalition may form can lead other groups to form competing coalitions and to struggle against each other. If stakes are high enough, such struggles can become violent. If we attribute the resulting violence to ethnic hatred, we may very well doubt that there is anything we can do. If we understand the way the violence arises from the situation, we may see more opportunities to end the conflict.

VIII

Summary. We very confidently attribute character traits to other people in order to explain their behaviour. But our attributions tend to be wildly incorrect and, in fact, there is no evidence that people differ in character traits. They differ in their situations and in their perceptions of their situations. They differ in their goals, strategies, neuoses, optimism, etc. But character traits do not explain what differences there are.

Our ordinary views about character traits can be explained without supposing that there are such traits. In trying to explain why someone has acted in a certain way, we concentrate on the figure and ignore the ground. We look at the agent and ignore the situation. We are naive in our understanding of the way others view a given situation. We suffer from a confirmation bias

that leads us to ignore evidence against our attributions of character.

It is very hard to do studies that might indicate whether or not people differ in character traits, but the few studies that have been done do not support this idea. We must conclude that, despite appearances, there is no empirical support for the existence of character traits.

Furthermore, it is clear that ordinary thinking about character traits has deplorable results, leading to massive misunderstanding of other people, promoting unnecessary hostility between individuals and groups, distorting discussions of law and public policy, and preventing the implementation of situational changes that could have useful results.

Department of Philosophy
Princeton University
Princeton
New Jersey 08544

Bibliography

Aristotle, 1985, *Nicomachean Ethics*, translated by T. Irwin. (Indianapolis, Indiana: Hackett).

Bennett, W. J., 1993, *The Book of Virtues*. (New York: Simon & Schuster).

Brandt, R. B., 1988, 'The structure of virtue'. *Midwest Studies in Philosophy XIII:* 64–82.

Darley, J. M. and Batson, C. D., 1973, '"From Jerusalem to Jericho": A Study of Situational and Dispositional Variables in Helping Behavior', *Journal of Personality and Social Psychology* 27.

Dawes, R. M., 1994, *House of Cards*. (New York: Free Press).

Doris, J. M. (forthcoming), *People Like Us: Personality and Moral Behavior*, (New York: Cambridge University Press).

Fischer, J. M. and Ravizza, M., editors, 1992, *Ethics: Problems and Principles*. (Fort Worth: Harcourt Brace Jovanovich).

Flanagan, O., 1991, *Varieties of Moral Personality*. (Cambridge, Mass.: Harvard University Press).

Hardin, R., 1995, *One for All: The Logic of Group Conflict*. (Princeton, NJ: Princeton University Press).

Harman, G., 1983, 'Human flourishing, ethics, and liberty', *Philosophy and Public Affairs* 12: 307–22.

Harman, G., 1996, 'Moral relativism,' part I of Harman & Thomson, 1996.

Harman, G. (forthcoming), 'Moral philosophy and linguistics', *Proceedings of the 20th World Congress of Philosophy* (Philosophy Documentation Center).

Harman, G. and Thomson, J. J., 1996, *Moral Relativism and Moral Objectivity*. (Oxford: Blackwell).

Hartshorne, H. and May, M. A., 1928, *Studies in the Nature of Character*, *I: Studies in Deceit*. (New York: Macmillan).

Holland, J. H., Holyoak, K. Jl, Nisbett, R. E. and Thagard, P. R., 1986, *Induction: Processes of inference, learning and discovery*. (Cambridge, Mass.: Bradford Books/M.I.T).
Hursthouse, R., 1996, 'Normative virtue ethics', in *How Should One Live? Essays on the Virtues*, ed. R. Crisp (Oxford: Oxford University Press), pp. 19–36.
Kagan, S., 1989, *The Limits of Morality*. (Oxford: Oxford University Press).
Laing, R. D., 1960, *The Divided Self*. (Chicago: Quadrangle).
Lewin, K., 1935, *Dynamic Theory of Personality*. (New York: McGraw-Hill).
Merritt, M. (forthcoming), 'Virtue Ethics and the Social Psychology of Character', Ph.D. Dissertation, University of California, Berkeley.
Millgram, S., 1963, 'Behavioral study of obedience.' *Journal of Abnormal and Social Psychology* 67.
McCloskey, M., 1983, 'Intuitive physics'. *Scientific American* 248: 122–30.
Nagel, T., 1979, 'Moral luck', in *Mortal Questions*. (Cambridge: Cambridge University Press).
Newcomb, T. M., 1929, *The Consistency of Certain Extrovert–Introvert Behavior Patterns in 51 Problem Boys*. (New York: Columbia University Teachers College Bureau of Publications).
Nisbett, R. E. and Ross, L., 1980, *Human Inference: Strategies and Shortcomings of Social Judgment*. (Englewood Cliffs, NJ: Prentice-Hall).
Railton, P., 1997, 'Made in the Shade: Moral Compatibilism and the Aims of Moral Theory', *Canadian Journal of Philosophy*, Supplementary Volume 21.
Ross, L., 1977, 'The intuitive psychologist and his shortcomings', in *Advances in Experimental Social Psychology*, Vol 10, ed. L. Berkowitz (New York: Academic Press).
Ross, L. and Nisbett, R., 1991, *The Person and the Situation: Perspectives of Social Psychology*. (New York: McGraw-Hill).
Skinner, B. F., 1974, *About Behaviorism*. (New York: Knopf).
Smith, A., 1759, 'Of the Influence of Fortune upon the Sentiments of Mankind, with regard to the Merit or Demerit of Actions'. *Theory of the Moral Sentiments*, Part II, Section III.
Taylor, R., 1991, *Virtue Ethics*. (Interlaken, New York: Linden Books).
Thomson, J. J., 1996, 'Evaluatives and directives', chapter 8 of Harman & Thomson, 1996.
Tversky, A. and Kahneman, D., 1974, 'Judgment under uncertainty: heuristics and biases,' *Science* 185: 1124-31.