**MANIPULATIVE ACTIONS: A CONCEPTUAL AND MORAL ANALYSIS**

**Robert Noggle**

**This is an author’s archived version of a paper published in *American Philosophical Quarterly,* Volume 33 (1996): 43-55. Any references should be to the final published version.**

**I**

A large portion of the wrongs that people commonly do to one another--especially to friends and loved ones--are forms of manipulation. Even ordinarily moral people who seldom violate rights to life, liberty, or property--people who would not assault, abduct, or steal from one another-- often engage in manipulation. Given the commonness of manipulation in everyday life, it is rather surprising that the philosophical literature on interpersonal manipulation is quite sparse. It is even more surprising in view of two recent trends in philosophical ethics. The first is the rise of interest among philosophers in the ethics of personal relationships, which has been sparked in large part by the development of feminist ethics. Second is the current popularity of Kantian ethics with its notion of treating persons as ends in themselves. An account of (and prohibition against) manipulative action should be a crucial part of an ethic of personal relationships and an ethic of treating persons as ends. Yet such accounts remain rather scarce.[[1]](#footnote-1)

One possible explanation of the sparseness of philosophical literature on manipulation might be that philosophers have assumed that manipulation always involves deception.[[2]](#footnote-2) If this were true, then we would not need a moral analysis of manipulation separate from that of deception: the wrongness of manipulation would be merely a function of the wrongness of deception. But there seem to be forms of manipulation that do not involve deception. Consider the following cases.

Case 1. Claiming to be a police officer trying to catch a dishonest bank teller, the swindler asks his victim to withdraw her life savings. He then absconds with the money.

Case 2. Charity scams often collect money for a good cause, but fail to mention that a very high percentage of the donations goes toward "overhead," including large salaries for the administrators and fund-raisers.

Case 3. In Shakespeare's *Othello*, Iago fills Othello's mind with what Othello takes to be evidence of Desdemona's infidelity, even though Iago knows that all of this "evidence" consists of red herrings.

Case 4. Iago plays upon Othello's anger and jealousy, so that when he becomes convinced that Desdemona has been unfaithful, he becomes enraged and murders her.

Case 5. Without telling any outright lies, an advertiser tries to create the impression that owning a certain car will make one the epitome of youth, vigor, and sex appeal.

Case 6. Satan tempts Christ who is fasting in the wilderness. He reminds him of his hunger and of the fact that he could turn the stones into bread.

Case 7. An adult sulks when he does not get his way. His friend, who already knew that he was upset, now gives in and gives him what he wants.

Case 8. B. F. Skinner's students supposedly conspired to pay attention to him only if he took a step toward the door. Soon--the story goes--they had conditioned Skinner to deliver his lecture from the hallway.

These all seem like cases of manipulation.[[3]](#footnote-3) Yet they are surely a diverse group. In addition to outright lies and other forms of deception, we have tempting, inciting, insinuating, conditioning, and playing on emotions. One wonders what, if anything, unites these cases in such a way that they are all instances of the same thing. That is one puzzle--a puzzle about the concept of manipulative action. A second puzzle--this one about the moral status of manipulative action--is that if deceptive manipulation is indeed wrong because it involves deception, then what makes *non-deceptive* manipulation wrong? I intend to offer an analysis which solves both puzzles, one which both captures the diversity of the phenomena to which the concept of manipulation applies, and which shows why all forms of manipulative action are morally blameworthy.

**II**

If we look closely at each of these cases, we notice that in many of them the victim has been led astray in some way. In the case of the swindlers this is obvious. Iago leads Othello so far astray that he eventually murders Desdemona. Satan tries to lead Christ astray by tempting him to eat when he should continue his fast. The advertiser leads the viewer of the commercial to mistakenly think the car will bring him youth, vigor, and sex appeal. I propose that we take this idea of leading the victim astray as the starting-point for an analysis of manipulative action.

The term `manipulation' suggests that the victim is treated as though she were some sort of object or machine. It's as though the manipulator controls his victim by "adjusting her psychological levers." There are three main "levers" which a manipulator can "operate." They are belief, desire, and emotion. This suggests that there are three main ways of manipulating someone, that is, three distinct ways that a manipulator can lead his victim astray. The paths from which the victim can be led astray are paths toward certain *ideals*. These are the ideals to which we strive to get our beliefs, desires, and emotions to conform. It is this striving that the manipulator attempts to thwart. To put the point a bit less metaphorically, there are certain norms or ideals which govern beliefs, desires, and emotions. I am suggesting that manipulative action is the attempt to get someone's beliefs, desires, or emotions to violate these norms, to fall short of these ideals.[[4]](#footnote-4)

Let's begin with deception, the attempt to operate the "lever of belief." I will use the term `direct deception' (`lying' in the vernacular) to include any assertion of a proposition that the asserter does not believe, with the intention of causing someone to believe that the proposition is true. When we engage in direct deception, we try to get someone to believe what we take to be false. Believing what is false is a failing: beliefs are "successful" if they are true.[[5]](#footnote-5) That is one of the *ideals* for beliefs. When we directly deceive, we try to get the victim to fall short of this ideal. So on the picture of manipulation I am suggesting, direct deception is a form of manipulative action. And that seems right.

But not all deception is direct, since not all of it involves assertion. Often a manipulator will try to call the victim's attention to red herrings, that is, to facts that are irrelevant to the victim's present situation. This is one of Iago's favorite tricks, and surely it counts as manipulation. Alternately, the manipulator can flood the victim's attention with so much irrelevant information that she is unable to concentrate on relevant information. Red herrings, then, can be used either to provide irrelevant inputs into the victim's reasoning or to crowd out relevant inputs so they cannot be used. The ideal for beliefs violated in both cases does not involve truth, for red herrings might be true. Rather, it is an ideal having to do with *relevance*. In particular, a belief ideally is attended to by the believer if and only if it is (true and) relevant to the situation at hand.[[6]](#footnote-6) So to call someone's attention to irrelevant beliefs is to violate this ideal and thus, on the analysis that I am suggesting, to act manipulatively. Again, this seems right.

Another indirect way to change someone's beliefs is to make insinuations, raise suspicions, or create impressions without actually lying, indeed, without even asserting anything. This can often involve imagery, the clever use of words with just the right connotations ("loaded language" or "code words"), leading questions, certain tones of voice, calling attention to beliefs one knows are false, and many other subtle techniques. These methods can set up patterns of salience among known facts, "set the agenda" for irrelevant lines of inquiry, lead the victim to misinterpret evidence, or reinforce false beliefs. Shakespeare's *Othello* is a study in this form of manipulation. Iago's insinuations about Desdemona and Cassio plant the seeds of suspicion in Othello's mind. Further insinuations lead Othello to take things as evidence for Desdemona's infidelity that he normally would not have thought twice about. (A misplaced handkerchief that turns up in Cassio's possession, for example.) Finally, he calls attention to those beliefs Othello has which support the contention that Desdemona has been unfaithful, even though he knows that those beliefs are false. The result, of course, is that without lying outright, Iago gets Othello to mistakenly believe that Desdemona has been unfaithful. This form of manipulation follows the pattern we have examined: an attempt to get the victim to violate the ideal of attending to all and only true and relevant beliefs. So the analysis I am suggesting counts them as manipulation. And this seems intuitively correct.

I now turn to the manipulation of desires. To a large extent, human motivation is instrumentally rational. That is, our desires usually conform to our beliefs about what we have reason to do (though often these beliefs are themselves irrational or false). But interwoven with this fabric of subjectively rational motivation we find pockets of subjectively *irrational* motivation. One source of such irrationality is psychological conditioning. Conditioning can apparently bypass the higher cognitive processes and produce motivation without necessarily changing any beliefs. Because it can do this, it can produce desires that do not conform to one's beliefs about what one has reason to do.[[7]](#footnote-7) (This is not to say, of course, that conditioning always produces irrational desires; more on this below.) For example it can make a professor desire to move toward the door without convincing him that it he has a reason to do so.[[8]](#footnote-8)

Desires can also be subjectively irrational if they display "motivational akrasia." Motivational akrasia occurs when one lacks sufficient motivation to do what one believes there is reason to do. This can occur in two distinct ways. First, a person can lack *any* desire to do what she believes there is reason to do. Or, more commonly, a person's strongest desire may not be for what she believes there is *most* reason to do, although she may have a desire (of some strength) to do each thing she believes she has *some* reason to do. This more common situation typically arises from most strongly desiring to do something that there is not most reason to do. Such situations are what we call temptation: though we may desire to perform the optimal action, and though we know that it is in fact what we have most reason to do, we have a stronger desire to perform some less optimal action.

I take it to be fairly obvious that a norm or ideal for desires is that they be subjectively rational, that they conform to one's beliefs about what there is most reason to do.[[9]](#footnote-9) It is then easy to see how the theory I am suggesting implies that various ways of influencing desires are manipulative. Conditioning someone in order to instill a desire that does not conform to the victim's beliefs is a manipulative action, for it is an attempt to instill a less than ideal desire. Likewise, to tempt someone is also to attempt to produce a situation in which one's motives are not ideal. According to both common sense and the analysis I am suggesting, to do either of these things is to act manipulatively.

The third main kind of manipulation involves the emotions. The ideals for emotions--the conditions under which each particular emotion is appropriate--vary from emotion to emotion.[[10]](#footnote-10) But in general, "positive" emotions such as joy, happiness, hope, and so on, are appropriate for someone who desires that *P* and believes that *P* is true or likely. Negative emotions such as anger, fear, and so forth, are appropriate for someone who desires that *P* and believes that *P* is false or unlikely. One norm or ideal for emotions is that the beliefs on which they are based should be true. Thus if I am happy that *P* but it turns out that not-*P*, then my happiness is not ideally appropriate. When Iago incites jealousy in Othello, he gets Othello to have an emotion that is inappropriate, since the belief that would make it appropriate, namely that Desdemona is unfaithful, is false. Again, the analysis I am suggesting agrees with what I take to be the commonsense judgment that such cases are cases of manipulative action.

The appropriateness of emotions depends on more than just the truth of the beliefs on which they are based. Consider jealousy, for example. It is appropriate only if the jealous person believes certain sorts of things, for example, that his lover is having an affair with someone else. Sadness, anger, annoyance, and shame are all negative emotions, and are all appropriate only in persons who want something to be the case that they think is not the case. Yet they are not all appropriate under the same conditions. This fact will turn out to have important consequences, as we shall see. But for the moment though, let's grant that we know what it means for an emotion to be appropriate. If a person tries to incite an inappropriate emotion in someone else, then on my account he acts manipulatively. Thus one person might try to incite another person to anger when only sadness or regret is appropriate. Consider the following story. Peter, Paul, and Mary have been planning to go on a picnic, but Mary is forced to cancel. Presumably some sadness or regret is appropriate here. Now unbeknownst to Peter, Paul dislikes Mary, and would like for Peter to do so as well. So Paul tries to get Peter to feel angry at Mary for canceling, rather than merely sad that she had to do so. On the analysis I am suggesting, Paul is acting manipulatively. For Paul attempts to incite an inappropriate--and thus a non-ideal--emotion. This judgment seems to me to accord well with common sense.

The manipulation of emotions can involve the attempt to violate yet another ideal for emotions, one which has to do with the fact that emotions make their objects salient. For this idea I draw on Ronald DeSousa's *The Rationality of Emotion*. DeSousa writes:

Emotions are species of determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies. . . . An emotion limits the range of information that the organism will take into account, the inferences actually drawn from a potential infinity, and the set of live options among which it will choose. . . . Emotions set the agenda for belief and desires: we might say they ask the questions that judgement answers with beliefs, and evaluate the prospects to which desire may or may not respond. As every committee chairman knows, questions have much to do with the determination of answers.[[11]](#footnote-11)

The idea here--and I think it is a correct and important one--is that emotions make important things conspicuous. Out of all the myriad bits of information which constantly bombard us, emotion helps determine what to pay attention to. Joy makes us pay attention to the fact that we've just matched all six lottery numbers. Fear makes us pay attention to the tornado coming our way. Love makes salient the fact that the person who just walked into the room is a loved one. When this system is working properly, emotion makes salient whatever is most important, most relevant to the situation at hand.[[12]](#footnote-12) It is a norm or ideal for emotions to function in this way.

Manipulators often attempt to violate this ideal. One way to do this is by sulking. Suppose Hansel and Gretel want to go to dinner together. Hansel wants to have Chinese food while Gretel wants Lebanese food. Suppose that Hansel realizes that the decision about where to eat will and should be a joint one (i.e., he has no special authority to decide for them both). And suppose that both Hansel and Gretel agree that taking turns is a fair and efficient way to make dining decisions. Suppose further that the last time they dined together they went where Hansel wanted to go. In spite of this, Hansel may decide to sulk in order to get Gretel to count his preferences as being more important than they are. Hansel thus may not set out to change the *content* of any of Gretel's beliefs; for he may know that Gretel is aware of his preferences. Rather his goal is to play on Gretel's sympathy in order to get her to assign greater *relevance* to his preference than it warrants in a time and context in which it is her preferences that should be most relevant. He is trying to make his preference a weightier factor in her deliberations than it ought to be. Specifically, Hansel wants Gretel to rate Hansel's preferences as being more relevant than her own preferences, more relevant than they are given that they have both agreed to take turns deciding where to dine and that they went where he wanted to go last time.

Or consider Jack, a manipulative person in a relationship. Suppose that his partner Jill is anxious about the possibility that her employer is considering a large lay-off. Suppose now that Jack and Jill are discussing something unrelated to her employment situation--whose family to spend the holidays with, for instance. It would be manipulative for Jack to remind Jill of her employment situation in order to induce anxiety that might allow him to bully her into agreeing to visit his family. For to do so is to play on her emotions in such a way that something irrelevant to the decision at hand becomes highly salient. Both Hansel and Jack attempt to "play on" another person's emotions in order get that person to make some fact more salient than it ought to be. And to intend for that to happen is to intend to move someone away from her ideal conditions with regard to emotions. Again, the analysis I am offering gets what I think is the intuitively correct result that such actions are manipulative.

**III**

So far I have been characterizing manipulative action as the attempt to "adjust someone's internal levers"--that is, someone's belief, emotion, or desire--away from their ideal settings. In some cases I have suggested that the ideal involves relevant information or appropriate emotion. In so doing, I have spoken as though the notions of relevance and appropriateness were relatively unproblematic. But this is far from the case. For we do not have decision procedures that tell us what is relevant (a fact that has bedeviled people working on artificial intelligence for some time[[13]](#footnote-13)). Nor do we have well-established criteria determining exactly what emotion is appropriate in every situation. Reasonable persons often disagree: is anger called for or mere annoyance, shame or mere regret, hope or worry? Indeed, it is not at all clear that there is any such thing as an "objective" standard for appropriate emotion or that there could be an algorithm for defining which information is relevant to which situation.

This fact leaves us with three options for characterizing manipulative action. First, we could appeal to objective ideals, hope that the problem of formulating them does not turn out to be intractable, and then either attempt to construct the appropriate ideal or wait for someone else to do so. This strategy does not seem promising. Second, we could relativize the ideal to what the person being influenced believes are the ideal settings for her. That is, a person would be manipulated if her psychological levers were adjusted away from what she--the victim--thought the ideal settings were. Third, we could relativize the ideal to what the *influencer* thinks are ideal settings *for the person being influenced*. I favor the third choice.

There are two reasons why I think that we should opt for this third choice and relativize the definition of manipulative action to what the actor takes to be the victim's ideal conditions. First, doing so preserves a conceptual parallel with lying. The notion of lying is defined relative to the beliefs of the liar. Thus I lie if I tell you something I do not believe, whether or not that thing is true. It seems obvious that the notion of directly deceptive manipulation should be parallel to, and perhaps even coextensive with, the notion of lying. Otherwise it could turn out that asserting something one takes to be both true and relevant could actually count as acting manipulatively, or that one could lie (even in a particularly sneaky way designed to get the victim to do something she would not otherwise have done) without acting manipulatively. Both of these results seem counterintuitive (and conceptually problematic). So it seems to me that we should define the concept of directly deceptive manipulation relative to the beliefs of the manipulator. That way, deceptive manipulation will turn out to be parallel to (and indeed coextensive with) lying. And if we relativize one form of manipulation to the beliefs of manipulator, then symmetry suggests that we should do so for the other forms as well.

Second, if we do not define manipulative action relative to the beliefs of the actor, then it would be possible--even common--to unknowingly perform manipulative actions. For an actor may have no idea what some other person takes to be his ideal conditions or what "*really*" is true, relevant, appropriate, and so forth. All an actor can be sure of is what she herself thinks would be ideal for the other person. It seems best to define manipulative action to reflect this fact. Then at least one person is always in a position to know whether a manipulative act has occurred, and that person is the actor. And this seems right, for on the analysis I am advocating, to act manipulatively is to *attempt* to lead someone astray. And who better to tell whether someone is attempting to do that than the person doing the attempting? If we want to promulgate norms prohibiting manipulative action, then we should define that notion in such a way that the actor can tell whether she is performing the action in question. For it makes little sense to prohibit someone from acting manipulatively if she is unable to tell when she is doing so.

**IV**

I am suggesting, in short, that we take manipulation to be an intentionally characterized action. That is, an action counts as manipulation if and only if the relevant intention is present. This allows us to place all the various kinds of manipulation under the same category. Moreover, it brings out what seems to be crucial to making an action manipulative: that it is done with a certain kind of insincere, conniving intention. If we define the notion of a manipulative action in terms of the intention of the actor, then it seems reasonable to use the point of view of the actor to define the ideal that the action is an attempt to violate. Returning to the leading astray metaphor, I propose that we classify all actions that are meant to lead astray--that is, all actions done with the *intention* to lead astray--as being of the same moral kind. If we do that, then for the purposes of deciding whether someone acts manipulatively, we need not worry about what the right path really is, or about what the victim thought it was, but about what the person doing the leading thought it was. And to do that, we will have to consult that person's beliefs about what the right path was.

In effect, I advocate drawing a conceptual parallel between acting manipulatively and lying. Lying bears the same relation to believing a false proposition as acting manipulatively bears to falling short of an ideal. To lie is to attempt to bring it about that someone believes what is false; to act manipulatively is the attempt to bring it about that someone falls short of the ideals. In both cases what is intended is crucial: if I try to get you to believe that P, when I think that not-P, then I lie--even if P turns out to be true. Similarly I act manipulatively if I try to get you to fall short of what I think is ideal, even if I am wrong.[[14]](#footnote-14)

This analysis has the following implications, all of which seem quite intuitive. Direct deception (as in Case 1) is a form of manipulative action, for it is an attempt to produce in someone a belief that is non-ideal because it is false. Indirect deception (as in Cases 2, 3, and 5) is a form of manipulation since it is also an attempt to get someone to violate the ideal of attending to all and only true relevant beliefs in their deliberations. Tempting someone (Case 6) is manipulative, since it is an attempt to get the person to have a non-ideal motivational state. Similarly the attempt to condition someone to have desires that do not conform to her beliefs (Case 8) is an attempt to get her to have desires that are non-ideal, and so it, too, is manipulative. The attempt to incite in someone else an inappropriate emotion (case 4) is manipulative, for such emotions fall short of the ideal that emotion should be appropriate. Similarly, the attempt to incite an emotion that assigns more salience to some fact than its relevance to the other person's situation warrants (Case 7) is manipulative, since such emotions also fall short of the ideal that emotions are to make things salient to the extent that they are relevant. It appears then, that the analysis of manipulative action as action done with the intent to lead astray can capture the fact that manipulative actions come in an otherwise bewildering array of forms.

**V**

The analysis I am proposing accords with the intuitive judgment that all of the eight cases with which I began are indeed cases of manipulative action. This is an important achievement for the analysis I am proposing. But just as important as what the analysis counts as manipulative action is what it does not count. In general, if I attempt to adjust your psychological levers but do not intend to move them *away from* their ideal settings, then on the analysis I am suggesting, I have not acted manipulatively toward you. What I have done is to engage in any of several forms of non-manipulative influence (some of which may be morally questionable for reasons not involving manipulation). On the account I am suggesting, the difference between non-manipulative influence and manipulation is that non-manipulative influence is not an attempt to produce a violation of the ideal conditions. Thus non-manipulative influence can involve the same kinds of influence that are used in manipulation.

Suppose you remind me of starving children in Rwanda, and describe their plight in vivid detail in order to get me to feel sad enough to assign (what you take to be) the morally proper relevance to their suffering. Surely you have not *manipulated* me, though you may have engaged in non-rational moral persuasion. Similarly, if one person tries to direct another person's attention to *relevant* beliefs, the first person is not manipulating the second; rather, she is offering a sort of counsel by pointing out what she thinks is pertinent information. Or if a psychologist uses conditioning to instill desires that conform to the patient's beliefs about what there is reason to do, then she is engaged in therapy rather than manipulation. These examples show that trying to move someone *toward* that person's ideal conditions is not in itself manipulative, even when it takes place by "non-rational" means. Rather it is what we might call "non-rational counselling." Apparently such counselling is often not immoral at all. In fact, in cases such as those involving behavior therapy or the appeal to sympathy in moral persuasion, it seems morally benign.

The fact that this analysis of manipulative action implies that some forms of non-rational influence are not manipulative is a very attractive feature. It is tempting to view non-rational influences as being inherently manipulative. Similarly, one might be tempted to define manipulation via some list of inherently manipulative influences. Though such views are tempting, I think that viewing certain kinds of influence as inherently manipulative is ultimately misguided. For this view cannot account for the fact that non-rational influences that in some contexts seem intuitively manipulative seem just as intuitively to be benign or even benevolent counselling in other settings. Depending on the direction I intend to move you relative to what I take to be the ideal for you, the very same activity could count as temptation or brainwashing, or as behavior therapy; as playing on the emotions or as moral persuasion; as throwing you off the trail with red herrings or as reminding you of important facts; as deception or as rational persuasion. What makes a form of influence manipulative is the *intent* of the person acting, in particular the direction in which she intends to move the other person's psychological levers. This is an important fact. And it is a fact which no analysis that depends on defining manipulation with respect to some list of inherently manipulative influences can capture.[[15]](#footnote-15)

The advantage of defining manipulative action relative to the intention of the actor is that doing so allows us to distinguish between manipulative and non-manipulative influences in a way that matches our intuitions and practices. For both our moral intuitions and our moral practices distinguish fairly sharply between such things as the sincere appeal to sympathy in moral persuasion and playing on the emotions, between brainwashing and behavior therapy, in short, between offering counsel and acting manipulatively. This is such an important achievement that I think we should take the base notion of manipulation to be acting manipulatively (as opposed, for instance, to being manipulated) and then define that notion relative to the intention of the actor. Doing so allows us to group together all the actions done with a certain kind of intention, and to separate from them seemingly similar actions which are done with very different intentions. In this way it allows us to draw a principled distinction between manipulation and counselling.

It is important to keep in mind that even though someone may engage in counselling rather than manipulation, other moral criticisms of his behavior may still be possible. For instance, we can criticize a counsellor for the way in which she offers her counsel. She may be pushy or cruel, or she may fail to respect our right to reject her counsel. We can also criticize the counsellor for thinking that certain things are true and relevant. For example a racist who attempts to incite racial fears may not intend to move the other person away from what *he*--mistakenly--takes to be the other person's ideal condition, and so we cannot accuse him of acting manipulatively. But of course we *can* criticize him for being a racist. And some beliefs, such as racist beliefs, are so indefensible and so pernicious that just having them can be highly immoral. So non-manipulative counselling by a sincere racist might even be worse than manipulation by a "pretend" racist *just because* the sincere racist really does hold those pernicious beliefs.

Thus we must be careful to distinguish between genuine manipulative action and merely behaving in ways that *appear* to be manipulative but which lack the relevant intention to lead astray. One might appear to act manipulatively if the influencer and the person being influenced disagree about what propositions are true and relevant to the latter's situation.[[16]](#footnote-16) Even if the influencer has a culpably false view of what is our ideal, the influence is not a manipulative action so long as it is sincere, that is, in accordance with what the influencer takes to be true, relevant, and appropriate. Thus the racist's view of what attitudes are reasonable to have is culpably false, but if he attempts to get us to have such attitudes, he does not attempt to manipulate. Similarly, some people who sulk in order to get others to pay more attention to them actually believe they are entitled to that attention. Probably this is the case with much of the sulking done by children. Often children (and some adults as well) have an inflated sense of their own importance; they genuinely believe that their pains and projects are (or ought to be) more significant than those of other people--not only to themselves but to others as well. Such cases are somewhat intricate morally: on my view such an agent does not in fact act manipulatively. Indeed, accusing her of acting manipulatively makes no more sense than accusing someone of lying when she says what he honestly, though incorrectly, believes.[[17]](#footnote-17) Yet, as we saw, such behavior may well be wrong on other grounds. Manipulation is a vice, but there are other vices as well; racism and selfishness are among them. And these vices call for action: racism calls out for refutation and education; selfishness calls out for attempts to elicit altruism and sympathy. But acknowledging this does not commit us to viewing these vices as the same vice as manipulation.

**VI**

On the present analysis of manipulative actions it will be difficult in practice to determine whether someone acts manipulatively. Consider a case where Smith gets Jones to be very angry about something. To decide whether Smith acts manipulatively toward Jones, we will have to know a number of things: how angry Smith thinks it is rational for Jones to be (which is a function of how much anger Smith thinks is appropriate for someone in Jones's present situation), how angry Smith thinks Jones actually is, and how Smith thinks Smith's actions will affect the level of Jones's anger. Obviously each of these is very difficult to figure out in practice. This suggests that if we want to discourage acts of manipulation we would be well advised to promulgate norms that prescribe guilt for one's own manipulative actions rather than to institute a system of blame or punishment for such acts. Because it is so difficult to tell when someone acts manipulatively, we will have to try and get people to look into their own hearts and examine their own intentions.[[18]](#footnote-18)

It is worth noting that in order to know whether some other person has acted manipulatively, one must first "get into the head" of that person and understand her beliefs, attitudes, and motives. This fact yields the rather paradoxical result that close personal relationships are the contexts in which we are most likely to know whether someone is acting manipulatively. And thus it is in those contexts in which blame and punishment would be most effective and fair. But in those contexts blame and punishment are very often less appropriate than other forms of moral persuasion. Thus it seems that because of difficulties inherent in the practical epistemology of manipulation that it is not a good candidate for blame and punishment. For one can seldom be sure of whether a stranger acts manipulatively, and there are usually better ways to convince friends and family not to act manipulatively than to engage in blame or punishment.

**VII**

I noted at the outset that manipulation cannot be wrong simply because it involves deception, since it does not always involve deception. Yet the two certainly seem to be related. Indeed, on the analysis I am suggesting, deception is a variety of manipulation. For that reason, it might seem that the wrongness of manipulation still might be related to the wrongness of deception. It is fairly commonly held that under normal circumstances we are under an implied contract to tell the truth. Thus for example, Charles Fried claims that use of the "institution of asserting" carries with it an implied promise to assert only what one believes.[[19]](#footnote-19) Our practices of asserting are built on the expectation that people will in general keep this promise. Indeed that expectation is built into the way we use language. Might we be under a similar contract not to manipulate each other?

Certainly we might *expect* (or demand) not to be manipulated, but this expectation (or demand) is not built into social life in the way that the expectation of truth-telling is built into the practice of asserting. A prohibition against manipulation is not a condition for the very possibility of social interaction in the way that a prohibition against lying is. Moreover, it is not clear that manipulation *could* be the subject of a contractual prohibition. For contracts have to be enforceable. And to enforce a contract, one needs to have some hope of knowing when it has been violated. Now we can often find out that someone has been lying since we can often attribute beliefs with reasonable accuracy and then compare what was said with what was believed. But the epistemic challenge of determining whether someone has performed other kinds of manipulative actions will generally be much more difficult to meet. For to do so, we will have to know a good deal more than whether the person who said that **P** believes that **P**; we will also have to know what mental states he attributes to his putative victim, what attitudes he thinks are "proper" for this other person, and how he thinks his activities will affect the other person's mental states. The problems with meeting this challenge are likely to be for all practical purposes intractable. So it seems that a contract not to manipulate would not be a viable option--either in a Rawlsian Original Position or for less hypothetical contractors.

But while there does not seem to be a general implicit contract not to manipulate, I do think that an implicit agreement to renounce manipulation is indeed embedded in lots of our special relationships. Indeed, the renunciation of manipulation would seem to be a precondition for genuine friendship, and for any other relationship which is based on mutual trust, respect, and a desire to promote the interests of the other person. Thus while an implicit contract with all of humanity may well be part of what makes lying to a stranger wrong, (non-deceptive) manipulation of strangers is not wrong on contract grounds. But it probably is wrong among friends and loved ones at least in part because of an implicit agreement. A renouncing of manipulation is probably as necessary to deep personal relationships as the renouncing of lying is to language. However this still does not give us a story about why manipulation is wrong in other contexts.

Given the diversity of phenomena that seem to count as manipulation, one might have expected that only a rather complex moral analysis could explain why manipulation qua manipulation is wrong. But in fact, I think that the story is fairly simple. I think that all cases of manipulative action involve a certain kind of intention--namely the intention to lead astray, to induce a violation of certain ideals. And I think that the reason it is wrong to do this is just that the conditions are in fact ideal, that to lead a person away from them is indeed to lead her *astray*.

The ideals that the manipulator attempts to get her victim to violate are bound up with goals of rational and moral agency. It is pretty clearly a goal of rational agents as such to attempt to have true beliefs. Another such goal is to deliberate from just that information which is relevant to the situation or deliberations at hand. So the goal of reasoning from ideal conditions would seem to be at the very heart of the notion of rationality. Just as important as the goal of reasoning from true relevant beliefs is the goal of acting on the results of one's reasoning. That is, it is a goal of rational agents as such to be motivated to do what one's deliberations show there is most reason to do. That goal is closely linked to the notion of agency. So the attempt to lead astray from these ideals of belief and desire is to attempt to thwart the goals of rational agency.

While the ideals for belief and desire are bound up with the goals of rational agency, the ideals for emotion seem more closely related to *moral* agency. Our emotions give flavor and texture to our lives as moral agents. They are not merely superfluous add-ons to rational calculators, but rather they turn us from cold calculators to warm beings with attachments and projects. To a very large extent they are what attach us to who and what we care about. Thus they are central to persons as beings with meaningful lives. And if they are not conceptually necessary to moral agency as such, it is pretty clear that they are necessary preconditions for *human* moral agency.

Acting manipulatively toward someone, then, is an affront to her as a rational and moral being, for it is an attempt to thwart her moral and rational agency which has as its goal the correct adjustment of her psychological levers. To attempt to thwart the goals someone has *qua* rational moral agent is pretty clearly to fail to respect her rational moral agency. And since a person's rational moral agency is crucial to her personhood, to fail to respect it is degrade her; it is to treat her as less than a person. And for that reason it is wrong[[20]](#footnote-20).

A pretty simple story, I admit. But I think that the fact that the story is so simple is actually a virtue of the analysis I am suggesting. For the fact that it is such a short step from the conceptual analysis of manipulation to its moral analysis is good reason to suppose we've gotten the conceptual analysis right. The notion of manipulation is, as a matter of common sense, closely tied to the idea of treating someone as a machine--as an object. We should expect that once we have an analysis of that notion, it should be pretty clear why acting manipulatively violates rational agency and fails to respect the personhood of the victim. And so we should not be surprised that it would be a short step from a conceptual analysis of acting manipulatively to a moral analysis[[21]](#footnote-21).

NOTES

1. In one of the few philosophical studies of manipulation in the literature ("Manipulation," *Ethics* 78 (1978): 338-47), Joel Rudinow develops a conception of manipulation that involves one person maneuvering another into a situation in which she has very limited options. I hesitate to call this manipulation; the term `maneuvering' (`railroading'?) seems more appropriate. In any case, it is not the kind of case I will examine here. In the cases Rudinow examines, the victim's circumstances, rather than the victim herself, are altered: one does not influence the victim from the inside, as it were, but rather from the outside. The difference between this and what I am calling manipulation is roughly the difference between herding or corralling sheep or cattle on the one hand, and riding a horse or driving a car on the other. [↑](#footnote-ref-1)
2. See Tom Beauchamp, "Manipulative Advertising," (in Tom Beauchamp and Norman Bowie, eds., *Ethical Theory and Business* [Englewood Cliffs, New Jersey: Prentice Hall, 1988], pp. 425ff.) for an explicit statement of what I suspect is a common view. [↑](#footnote-ref-2)
3. Those who do not share this (largely linguistic) intuition can take this project as an argument that something unites these cases, and can treat `manipulation' as a technical term for whatever that might be. Alternately, they could construe what I am doing as recommending a certain way of using the moral term `manipulation'. Whether or not one should accept such a recommendation depends in part on whether the suggested use of the term is useful. One thing counting toward the usefulness of a moral term is the extent to which it makes salient moral features of its referent that may not have been salient before. I take it that if I am successful in showing that the above cases share certain moral features, then I will have done this. [↑](#footnote-ref-3)
4. Strictly speaking, I want to suggest that to act manipulatively is to attempt to get someone's desires, beliefs, or emotions to violate the norms *more than they already do*, or to fall *further* short of the ideals than they already do. Alas, few of us have ideal beliefs, desires, or emotions. [↑](#footnote-ref-4)
5. The notion of true beliefs being "successful" if they are true goes back to G.E.M. Anscombe's *Intention* (Ithaca: Cornell University Press, 1969). For a survey of recent literature on the notion of success conditions for mental states, and an attempt to spell out the notion, see Lloyd Humberstone, "Direction of Fit" (*Mind* 101 [1992]: 59-83). [↑](#footnote-ref-5)
6. `Situation' here should be construed widely enough to include such contexts as those involving deliberation about some future situation. Obviously, the notion of relevance is somewhat problematic; some of these problems will be of theoretical importance. More on this below. The "if and only if" here means that indirect deception by (intentional) omission of relevant information (such as we have in the charity scam in Case 2) counts as manipulative action. [↑](#footnote-ref-6)
7. I discuss in detail this ability of conditioning to inculcate motivation without changing beliefs--and the moral implications of this ability--in "Autonomy, Value, and Conditioned Desire," *American Philosophical Quarterly* 32 (1995): 57-69. [↑](#footnote-ref-7)
8. This analysis of the example depends on the assumption that the process really is one of conditioning, and that the professor does not simply come to believe that his lecture is more comprehensible or interesting the closer he is to the door. [↑](#footnote-ref-8)
9. I take it that this formulation of the ideal conditions for desire includes, roughly, Davidson's claim that one's strongest desire should be for what one has most reason to do *on balance* (see Davidson, "How Is Weakness of the Will Possible?" in Joel Feinberg, ed., *Moral Concepts* [Oxford: Clarendon Press, 1970], reprinted in Davidson, ed., *Essays on Actions and Events* [Oxford: Clarendon Press, 1980]). Motivational akrasia or subjectively irrational desires are phenomena in which motivation gets out of line with what one has the best on balance reason to do. [↑](#footnote-ref-9)
10. For much more in-depth treatments of the rationality of emotions, see Ronald DeSousa, *The Rationality of Emotion* (Cambridge: MIT Press, 1987), Robert Gordon, *The Structure of Emotions* (Cambridge: Cambridge University Press, 1987), and Patricia Greenspan, *Emotions and Reasons* (New York: Routledge, 1988). [↑](#footnote-ref-10)
11. Ronald DeSousa, *The Rationality of Emotion* (Cambridge, Mass.: MIT Press, 1987), pp. 195-6. [↑](#footnote-ref-11)
12. None of this, of course, is to deny that emotions do other things. Indeed, somewhat paradoxically, it is largely the function of emotion to make certain things relevant to us; that is, what we care about, what is important to us, is often a function of complexes of emotions. Yet this does not prevent emotions from also pointing out--and thus reinforcing--the fact that we care about certain things. [↑](#footnote-ref-12)
13. The problem of defining relevance is essentially the so-called frame problem. For more on that, see Jerry Fodor, *The Modularity of Mind* (Cambridge, Mass.: MIT Press, 1983, pp. 112ff.). [↑](#footnote-ref-13)
14. This way of thinking about manipulative action owes much to the work of Claudia Mills, work I have found very useful in helping to frame the issues I deal with here. My characterization of manipulative action is in many ways similar to one she offers in "Politics and Manipulation," (*Social Theory and Practice* 21: 97-112). Yet there are significant differences as well. Whereas I characterize manipulation as the attempt to move beliefs, desires, or emotions away from some ideal, she characterizes it as the attempt to use bad reasons to persuade. Both analyses are relativized the same way, hers to what the manipulator thinks are bad reasons, mine to what the manipulator thinks are the ideals for the victim's beliefs, desires, and emotions. Thus the two analyses are parallel, or nearly so, with regard to the manipulation of belief. The main difference is that Mills's account does not deal with, and her analysis does not handle, some cases of manipulation that involve changing of desires and emotions rather than persuasion. In such cases the manipulator does not attempt to persuade anyone of anything, with bad reasons or otherwise. Indeed, part of what is frustrating about such cases is that the manipulator has gone outside the institution of offering reasons altogether. [↑](#footnote-ref-14)
15. And it will not help to define manipulative actions as those that occur "behind the back," that is, as involving processes of which the victim is not aware. For I may be quite aware of being tempted or of having my emotions played upon, and yet the actions may still be manipulative--and successful. [↑](#footnote-ref-15)
16. Of course it might also happen that A acts manipulatively toward B but B does not recognize A's action for what it is because A and B disagree about what is relevant to the situation at hand. To adapt an example suggested by an anonymous referee, suppose A and B are vying for the affection of C. A, a noted chemist, has the habit of ruminating on his work over dinner--a habit which annoys C. At a dinner attended by A, B, and C, the devious A mentions some new finding relevant to B's work, knowing that doing so will trigger a long rumination by A that will annoy C. B might be very grateful to A for the "relevant" information, yet, on my account, A acts manipulatively because she intends do distract B from (what A sees as) the situation at hand for B. [↑](#footnote-ref-16)
17. Some people find this counterintuitive. For example, Gillian Brock has suggested to me the case of a sincere advertiser engaged in non-rational persuasion as a counterexample to my view. For my part, I think that the very sincerity of such a person makes this a case of non-manipulative influence rather than manipulative action. (Of course we may think certain kinds of advertizing practices are immoral even if they are not manipulative; the fact that the advertizer is sincere does not necessarily make his actions morally legitimate even if it does make them non-manipulative.) But just as the eccentric who sincerely preaches the flatness of the earth does not lie, so the sincere advertiser (one is tempted to say: if any such there be) does not act manipulatively. Of course we might find it as eccentric to believe certain claims in advertisements as it would be to believe that the earth is flat, but that is another matter entirely. [↑](#footnote-ref-17)
18. It is probably in part because of these epistemic difficulties that we don't have laws against any kind of manipulation other than certain forms of direct deception (like fraud and perjury) though we do have them against such means of immoral influence as coercion and the use of brute force. [↑](#footnote-ref-18)
19. Cf. Charles Fried, *Right and Wrong* [Cambridge, Mass.: Harvard University Press, 1978], pp. 54-80. Contrast Thomas E. Hill, Jr., "Autonomy and Benevolent Lies" [in *Autonomy and Self-Respect* [Cambridge, Mass.: Cambridge University Press, 1991] for some skepticism about an implied truth-telling contract. [↑](#footnote-ref-19)
20. It is certainly coherent to maintain (and indeed, I want to maintain) that there are worse things than being manipulated. Manipulative action, on my account, is a form of degradation. But while the fact that an action degrades someone always counts against it, it may not always do so in a decisive way. One interestingly complex case occurs when A acts manipulatively toward B to bring about some non-ideal state *from which* she intends to move B to an ideal state. Such benevolent or paternalistic instrumental manipulation might sometimes be the only way to get a person into some ideal state. Whether such cases are morally permissible will depend on, inter alia, the importance of this otherwise unattainable final state and the extent to which the intermediate state falls short of the ideal. A detailed account of when manipulative actions--or other actions that degrade their victims--are permissible would require a general theory of the interplay of deontological and consequentialist considerations. Even if I had such a theory, there would be no room to present it here. [↑](#footnote-ref-20)
21. . Thanks are due to Gerald Dworkin, Charles Mills, Peter Hylton, Claudia Mills, Gillian Brock, and an anonymous reader for American Philosophical Quarterly for useful suggestions and comments on earlier drafts, as well as to Yvonne Micheletti, Lon Becker, Jim Harrington, Mike Kearns, and Josh Gert for useful feedback at an informal presentation of the paper. A shortened version of this paper was presented at the Pacific Division Meeting of the American Philosophical Association in San Francisco, March 31, 1995. Thanks are due to my commentator, Thomas Buller, for his insightful remarks, to the audience for their useful feedback, and to the APA for a Graduate Student Essay Award which covered my travel expenses. Finally, I thank the Graduate College at the University of Illinois at Chicago for a fellowship which supported work on this paper. [↑](#footnote-ref-21)