MANIPULATION, MISTAKES, AND THE QUESTION OF PERSPECTIVE

Word Count: 2998 (not counting title, notes, references, and abstract)

Abstract: A promising account of interpersonal manipulation characterizes it as inducing another person to adopt a mistaken or faulty mental state. Thus, Iago manipulates Othello by making him feel anger and jealousy that are inappropriate to the situation. This account faces a question: Whose perspective determines whether a mental state is mistaken for the purposes of defining manipulation? Is Iago’s behavior manipulative because Iago believes that Othello has no reason to be jealous, or because there really is no reason for Othello to be jealous? This paper seeks to answer this question by exploring the parallels between manipulation and deception. It concludes that: (a) acting manipulatively involves trying to get someone to adopt what the manipulator regards as a mistaken mental state, but (b) a person has been manipulated only when manipulative action induces her to adopt a mental state that really is mistaken.

Consider these cases of everyday interpersonal manipulation:

- Iago plays on Othello’s emotions to work him into a jealous rage.
- Manny dissuades Tara from leaving their small town for college in the big city by over-emphasizing college costs and sensationalizing the city’s crime rate, while dismissing and downplaying the benefits a college education would give her.
- Abusive Al gaslights his partner Patricia into doubting her own (initially accurate) judgement about how toxic their relationship is.
- A parent guilt-trips his child about her choice of college by repeatedly asking, “After all I’ve done for you, are you really going to break my heart by not attending my alma mater?”

A promising account of manipulation, which I will call the Mistake Account, characterizes manipulation as inducing a person to adopt a mistaken mental state. On this account, manipulation resembles lying and deception: Whereas lies and deception involve inducing someone to adopt a mistaken belief, manipulation involves inducing a person to adopt any sort of mistaken mental state—belief, desire, emotion, judgment, etc.

The Mistake Account captures the intuition that the cases above all involve manipulation: Iago’s behavior is manipulative because it induces Othello to feel emotions of jealousy and anger which are not justified by the facts. Al’s gaslighting is manipulative because it gets Patricia to mistakenly doubt her own judgement. Manny’s behavior is manipulative because it distorts Tara’s
judgement by getting her to pay too much attention to certain things and too little attention to others. The parent’s guilt trip is manipulative because it gets the child to feel inappropriately guilty over her choice of college.

Moreover, the Mistake Account recognizes that whether a kind of influence seems manipulative depends on how it is used. Iago appeals to Othello’s emotions, but emotional appeals—such as the appeal to empathy in moral persuasion—do not seem manipulative when the emotion is appropriate or warranted. Emphasizing certain factors over others does not seem manipulative when it helps bring important factors into proper focus. Getting an intoxicated person to doubt her overconfident judgement that she can drive home safely hardly seems like manipulative gaslighting. Nor does it seem manipulative to get someone to feel guilty about a seriously immoral action.

Despite these advantages, the Mistake Account faces what I will call the **Perspective Question:** Whose perspective determines whether a mental state is mistaken for the purposes of determining whether manipulation has occurred? Should we say that a person is manipulated only if she adopts a mental state that *really is* mistaken? Or should we say that a person is manipulated only if she adopts a mental state that is *believed by the manipulator* to be mistaken? Is Iago’s behavior manipulative because *Iago believes* that Othello has no reason to be jealous, or because there *really is* no reason for Othello to be jealous? Both answers have their proponents.

The Perspective Question is not always pressing. Suppose that influencer N gets another person to adopt mental state S. Matters are simple when S is *both* objectively mistaken and believed by N to be mistaken. The Mistake Account regards such situations as clear cases of manipulation. Matters are also simple when S is *neither* objectively mistaken nor believed by N to be mistaken. The Mistake Account finds no manipulation in such cases.

But what happens when N regards S as mistaken when, objectively, it is not? Or when S is objectively mistaken, but N doesn’t know that? Are such cases manipulation? Consider this example of the first kind of case:
Billy and Todd are co-workers. Billy bullied Todd severely as a child and remains contemptuous of Todd as an adult. But Billy hides his contempt from Todd, pretending to be Todd’s friend and ally while ridiculing and undermining Todd behind his back. Through a combination of naïveté and self-deception, Todd regards Billy as a friend, and dismisses his childhood bullying as kids being kids. Irene, a new co-worker, has just met Billy and Todd. When Todd tells her what a good friend Billy is, she believes him. However, she sees their friendship as an obstacle to her own advancement, and she resolves to drive a wedge between them. She repeatedly draws Todd’s attention to what both she and Todd initially regard as Billy’s good-natured ribbing, but she insinuates that Billy’s behavior is a sign of contempt for Todd. Eventually, her efforts lead Todd to become suspicious, resentful, and then angry at Billy, until he finally regards Billy as a bully rather than a friend. Unbeknownst to Irene, Todd’s new attitudes are far more appropriate than those they replaced.

Now consider this example of the second kind of case:

Aunt Vicky is a vaccine skeptic. She reads anti-vaccine propaganda as well as scientifically sound information about vaccines, and she has reached no definitive conclusion about vaccine safety. However, she has become so fixated on the rare adverse events associated with vaccines that she over-estimates the threat they pose. Lacking experience with the serious diseases that vaccines prevent, she underestimates the risks of vaccine-refusal. Despite having drawn no definitive conclusion about vaccine safety, she has developed an irrational—but sincere—fear of vaccination. Without making any assertions about vaccine safety, Vicky encourages others to share her fears about them. Her encouragement causes Cousin Carl to become inappropriately fearful of vaccines.5

Are these cases of manipulation? Initial intuitions will likely vary. My goal is to work out what the Mistake Account should say about such cases. If that account’s theoretical virtues are sufficiently compelling, then perhaps we will be willing to abandon any intuitions that it cannot accommodate, especially if they are weak and/or not widespread, or if they can be explained away.
My approach will be to build on the fact that the Mistake Account treats:

(1) Mandy manipulated Tim.

As being very similar to

(2) Larry lied to Tim.

and

(3) Dora deceived Tim.

First, we must decide whether (1) is more like (2) or more like (3).

One important difference between them is that ‘deceive’ is a success term and ‘lie’ is not. Suppose that Larry and Dora both believe that there is candy in the cupboard. However, they do not want Tim to eat the candy, so they try to make him believe that the cupboard is devoid of candy. Larry flat-out lies, while Dora resorts to “mere” deception. Suppose that their efforts fail. It would be natural, if slightly stilted, to say:

(4) Larry lied to Tim about the cupboard’s contents, but Tim nevertheless continued to believe that it contains candy.

Less stiltedly, we might say:

(5) Larry lied to Tim, but Tim didn’t fall for it.

But it would be odd to say:

*(6) Dora deceived Tim, but he didn’t fall for it.

Instead, we would say:

(7) Dora tried to deceive Tim, but he didn’t fall for it.

In short, I cannot be said to have deceived you unless I succeeded in causing you to believe something false. But I can be said to have lied to you even if I did not succeed in causing you to believe something false.6

What about manipulation? Suppose that Mandy wants to get Tim to stay home from a party by making him afraid to drive because it is snowing lightly. Mandy exaggerates the amount of snow that is falling, and recounts graphic tales of accidents on slippery roads. However, Mandy does not succeed in making Tim too afraid to drive to the party. We would not say:

*(8) Mandy manipulated Tim, but he didn’t fall for it.
Instead, we would say:

(9) Mandy tried to manipulate Tim, but he didn’t fall for it.

So ‘manipulate,’ seems to be a success term like ‘deceive.’ This finding is hardly earth-shattering on its own. But pursuing further the parallels between ‘manipulate’ and ‘deceive’ might provide a principled answer to the Perspective Question.

Suppose that, unbeknownst to Dora, Larry ate the candy from the cupboard. Still wanting to keep Tim from eating the candy that she (mistakenly) believes is there, Dora tries to get Tom to believe that the cupboard lacks candy. Suppose that Dora succeeds in this. However, since there was no candy in the cupboard, it would seem odd to say:

*(10) Dora deceived Tom into believing that there was no candy in the cupboard.

The reason this sounds odd, of course, is that what Dora got Tim to believe was, objectively, true. More generally, it seems odd to say that one person deceived another into believing that P when P turns out to be true. Thus, it seems to be a necessary condition for being deceived that one come to believe something that really is false. Hence, there appears to be an objective standard for being deceived.

Of course, our reluctance to say that Tim was deceived when the belief he acquired was true is not a moral free pass for Dora. Even if she did not deceive Tom, she tried to do so. We might say that she acted (or behaved) deceptively toward Tim. This accusation is based on Dora’s belief about the truth-value of the belief she tried to get Tim to adopt, and not its actual truth-value. Thus, we seem to employ a subjective standard for whether an agent acted or behaved deceptively.

Apparently, then, we have an asymmetry: To determine whether a person has been deceived, we employ an objective standard, asking whether the relevant proposition really was false. But to determine whether a person acted deceptively, we employ a subjective standard, asking whether the influencer believed that the relevant proposition was false.
This asymmetry seems puzzling until we see that it reflects two related but distinct concerns about deception. One is the badness of having been tricked or fooled into acquiring a false belief. Since belief aims at truth, believing falsely is a kind of failure. It is a failure in and of itself, but it also often leads to other misfortunes for the person who believes what is false. We needn’t decide here whether false belief is intrinsically or merely instrumentally unfortunate. It is enough to note that we value having true beliefs, and that we regard having a false belief as a misfortune. When this misfortune is the deliberate work of another person, it is of special concern. There is something especially bad about having acquired a false belief through the deliberate action of a deceiver. The misfortune of believing falsely is compounded by the insult of having been made a fool of. Now both of these bad things occur only when the person has been gotten to believe something that really is false, so this concern about what happens to the target of deception only applies when the person is gotten to believe something objectively false. Thus, the objective standard for having been deceived reflects our concern with what has happened to the victim of successful deception. It is a bad thing that only happens to a person only when she is tricked into adopting an objectively false belief.

A second concern about deception is the behavior of the person doing the deceiving. If being deceived is a misfortune, then it is blameworthy to attempt to foist that misfortune upon another person. This blame attaches not only when the deception is successful; it also applies to attempts to deceive. When we censure someone for trying to deceive, or acting deceptively, our focus is the agent’s morally improper intention, rather than what happened to the patient. What is morally blameworthy is intending to induce someone to adopt a false belief. But whether this morally inappropriate intention exists depends not on whether the belief in question really was false, but on whether the putative deceiver believed it to be false. That is, the moral impropriety of the intention depends on what the agent took to be the truth. Consequently, acting deceptively is a morally blameworthy behavior that occurs when the influencer intends for the target to acquire what the influencer takes to be a false belief.
It seems reasonable to have a similar pair of related but distinct concerns about manipulation: We care about the misfortune of being tricked into making a mistake in how we feel, what we pay attention to, our confidence in our own judgments, etc. This concern is with what happens to the target of manipulation, and it applies only when the target really does make a mistake. We are also concerned with the bad intention of a person who tries to foist a mistake on another person. But what makes that intention bad is not whether the mental state really was a mistake, but whether the putative manipulator thought it was a mistake.

This parallel between deception and manipulation suggests that we ought to employ an objective standard for *being manipulated* and a subjective standard for *acting manipulatively*. Accordingly, we should not say that Todd was manipulated. After all, despite Irene’s efforts, Todd made no mistake. Objectively speaking, Todd’s attitude toward Billy became less mistaken, and if the Mistake Account identifies manipulation with being induced to make a mistake, then it makes sense to deny that Todd has suffered the misfortune of having been manipulated.

But this does not exonerate Irene. Her attempt to foist a mistake upon Todd is subject to moral censure even though—unbeknownst to her—it resulted in Todd having an objectively more appropriate attitude toward Billy. Her intent was to induce Todd to make a mistake. The fact that she was wrong about the appropriateness of the feelings she tried to induce in Todd does not make her any less blameworthy for acting with bad intent. The fact that Todd did not suffer the misfortune of being tricked into making a mistake does not exonerate Irene for trying to foist that misfortune upon him.9

What about Vicky? Certainly, her conduct is far from ideal. Her critical thinking is—perhaps culpably—deficient, and she acts recklessly in spreading her unjustified fears. But the fact that Vicky’s behavior is deficient in some respects does not entail that it is deficient in every respect. In particular, Vicky’s intention was *not* to foist a mistaken fear on Carl, for she does not see this fear as being mistaken. Vicky is sincere in her fear, and well-meaning in her attempts to get others to share it. And if we adopt the suggestion here, we should regard her sincerity—misguided as it is—as a defense against the charge that she acted manipulatively.
Was Carl manipulated? This is trickier. If *being deceived* names the misfortune of being successfully tricked into a false belief by someone who is deliberately trying to foist a false belief upon one, and if we maintain the parallel between deception and manipulation, then we should not say that Carl was manipulated. The fear that Carl adopted was mistaken, but the behavior on the part of Vicky that induced it was not manipulative, for she did not intend to induce him to make a mistake. Vicky was sincere rather than conniving. These reflections suggest that it would be best to say that Carl was mistaken but not manipulated.

Nevertheless, some people still might want to say that Carl was manipulated. I concede that it would be a stretch to claim that saying that Carl was manipulated is just as clearly incorrect as saying that “Dora manipulated Tim, but Tim did not fall for it.” Perhaps the parallel with deception can help us decide what to say here. Suppose that Henry tells Tim that there is candy in the cupboard, and that this is what Henry sincerely but mistakenly believes. On this basis, Tim acquires the false belief that there is candy in the cupboard. Although I think it better not to say that Tim was deceived, perhaps saying so is not grossly incorrect, given that Henry did, in fact, cause him to acquire a false belief. Among philosophers, the view that one is only deceived if one is *deliberately* gotten to believe what is false seems dominant, but it is not universal.\(^ \text{10} \)

Moreover, in ordinary discourse, we sometimes speak of a person being deceived by an inanimate object—which suggests that intent is not necessary for deception. It is neither incomprehensible nor obviously wrong to say that a driver was deceived by black ice. But perhaps we can treat such usage as metaphorical: In certain respects, it is *as though* the black ice was trying to deceive the driver. Perhaps we should say a similar thing about unintentional or accidental deception: Insofar as Henry caused Tim to acquire a false belief, in certain respects it is *as though* Henry was trying to deceive him. Henry might say he was ‘deceived’ to lament his false belief, just as a driver might lament his false belief by claiming that the black ice deceived him. However, such usage, though comprehensible, is potentially misleading: While it is true that Tim suffered the misfortune of a mistaken belief, he did not also suffer the indignity of having someone deliberately trick him. Consequently, anyone who feels compelled to describe Tim as having been deceived in this case would do well to say that he was accidentally or unintentionally deceived.
Similarly, we could treat the claim that Carl was manipulated as metaphorical: After all, insofar as Vicky’s influence caused him to acquire an irrational fear, in certain respects it was as though Vicky had manipulated him. Here again, although such usage may be comprehensible, and even tempting, it is potentially misleading because it suggests the false claim that Vicky intended to trick Carl into a mistaken fear. Carl may have suffered the badness of acquiring a mistaken fear, but he did not suffer the indignity of someone deliberately tricking him. Consequently, anyone who feels compelled to describe Carl as having been manipulated would do well to say that he was accidentally or unintentionally manipulated. Otherwise, one might insinuate that Vicky had a blameworthy intention that she did not have. That insinuation might itself count as—perhaps unintentional—manipulation.

Notes

1 The qualifications “everyday” and “interpersonal” are meant to distinguish the sort of manipulation to be discussed here from the forms of manipulation often discussed in the free will literature where, for instance, a person has her entire psychology engineered by some extra-ordinary (and perhaps supernatural) agency.

2 It is also sometimes called the “Trickery Account” (Noggle 2018).

3 For an account of gaslighting similar to the one given here, as well as a helpful discussion of alternative uses of the term ‘gaslighting,’ see (Stark 2019).

4 Claudia Mill and Robert Noggle have argued that manipulation occurs when the influencer attempts to induce the target to adopt what the influencer regards as a mistaken mental state (Mills 1995; Noggle 1996). By contrast, Jason Hanna argues that we should define manipulation in terms of the attempt to induce the target to adopt an objectively mistaken mental state (Hanna 2015, 634; Sunstein 2016, 89). Anne Barnhill suggests that our usage of the term ‘manipulation’ is inconsistent on the question of whose standards determine whether the influencer attempts to induce the target to adopt a mistaken mental state (Barnhill 2014). In theory, we could also define “mistake” in terms of the target’s beliefs, so that Tom is manipulated only if Tom acquires what Tom regards (or, perhaps, what Tom would regard, on reflection) as a mistaken mental state. To the best of my knowledge, however, no one has defended this view.

5 Readers who do not regard deception as a form of manipulation, and who are worried that this case is more like the former than the latter should imagine that Vicky is very careful not to try to get anyone to believe that vaccines are unsafe, and that Cousin Carl was agnostic about vaccine safety both before and after interacting with Vicky. If one is tempted to object that it makes no sense to see Carl as adopting an irrational fear without also adopting a false belief, I would simply point out that belief and fear are not so tightly related as to preclude fearing a thing without having a definite belief that that thing is dangerous. I have many such fears, and chances are that the reader does as well.

6 In fact, lying is often defined in terms of an intent to deceive. (However, there are non-standard cases where the connection between lying and attempting to deceive is tenuous at best, and arguably absent altogether. Such cases lead some revisionists to suggest that the intent to deceive is not a necessary condition for lying at all (Mahon 2016).
Fortunately, the complex arguments about this matter need not detain us here, since I will be using deception rather than lying as a model for understanding manipulation.)

7 This is not to deny that a person can be deceived into believing some false statement from which he then goes on to derive some true statement.

8 The distinction here is similar to the distinction that Bernard Williams draws between the virtues of accuracy and sincerity in *Truth and Truthfulness* (Williams 2004). The organizing principle of that book is that we have distinct values and concerns with regard to truth. Very roughly, “accuracy” is the name he gives to our concern with having our beliefs be true, and “sincerity” is the name that he gives to our concern that people avoid asserting what they believe to be false.

9 A separate question, which I will not address here, is whether Irene’s blameworthiness is diminished by the failure of her attempt. The answer to this question will depend on whether, in general, where it is blameworthy to attempt to do A, it is less blameworthy to make a failed attempt to do A. Answering this question is beyond the scope of the present paper.

10 See (Mahon 2016, 40–41) for discussion and references.

**References**


