PHIL 640: Honesty

Brian Weatherson

November 15, 2021

Priority Questions

We describe both people and actions as honest. One interesting philosophical question concerns priority. Both of these things are true, but which of them are explanatory?

1. Honest actions are actions that honest people would be disposed to do.
2. Honest people are people who are disposed to do honest actions.

Miller says that 1 is explanatory and 2 is not; what makes something an honest action is that it is something an honest person would do. But I didn't really see an argument for this, and I could see a case made for the other side.

Indeed, I wasn't sure that Miller's own theory supported this direction of explanation. It's easy to at least gesture at what makes an action honest. Something about being truthful, or pro-truth, or something like that. And then you can gesture at what makes a person honest - they do those kinds of actions. (Perhaps typically, or perhaps in the important cases.) What's the story in the reverse direction?

Anyway, this is (I think) what he means when he says the virtue is prior to the actions. And I don't know what the argument is for this priority claim, even though it seems like a fairly important point to me.

Honest Actions

Miller gives an interesting list of five kinds of behavior that he thinks involve honesty. And we should be interested in (at least) three questions about this list.

1. Is it a correct description of how the term 'honest' is used in English?
2. Which of these kinds are such that *every* instance of them is an instance of dishonesty, and which of them are such that *typical* instances of them are instances of dishonesty?
3. What do these things have in common?

That said, here is the list.

- Lying
- Stealing
- Cheating
- Promise Breaking
- Deceiving

My view is that the first and the fifth are clearly instances of dishonesty, and the third might be - depending on what you mean by cheating. But I'm really not sure that stealing or promise breaking are. We'll come back to this when we look at Miller's explanation of why they are dishonest.

On cheating, it's perhaps stepping back to think about what we call cheating. Not all kinds of rule violations are cheating. A football player (in any code of football) who is offside isn't cheating, even if they break a rule. Someone who steps out of bounds while corralling a ball isn't cheating. What we call cheating, as opposed to mere rule-violation, I think involves some element of deception. So it isn't really a separate category.

That said, I'm not sure the later discussion of cheating, when Miller discusses steroids in baseball, really gets to the heart of the matter. I don't think the steroids taker is representing themselves as not using steroids. (Ignore cases where they expressly say they aren't using steroids - those obviously are cases of deception!) Imagine a league where this kind of cheating is rampant. (As it was in baseball until recently!) It seems weird to say that everyone is implicitly misrepresenting their status. The kind of representation Miller has in mind here only works if the typical baseball player is not cheating in this way.

There are also hard questions about what it takes for something to be 'natural'. Does someone who uses a hyperbaric chamber to speed recovery from injury count as using natural methods? (It might be relevant here to note that one of the main benefits of steroids in baseball just is speeding recovery time.) I'm not sure how steroids are unnatural but the hyperbaric chamber is natural. It's just that we have a rule against one but not the other.

So here are three questions I'd like us to have on the table. (And if this was a more general philosophy of sport class we'd spend way longer on each.)

1. What is the difference between cheating and rule-breaking?
2. Must cheating, as opposed to rule-breaking, involve deception?
3. Must cheating involve misrepresentation? (Subsidiary question: what facts were the Houston Astros distorting/misrepresenting when they cheated?)

The Honest Thief

We're not going to go over the plot of either the Dostoyevsky story or the Liam Neeson movie - though if someone wants to discuss either we can. (Though preferably with warning - I haven't read/seen either!) But I do think it's interesting that we do have the phrase 'honest thief' in English (and I guess an equivalent in Russian). I'll come back to what Miller says about this case, but I wanted to flag two questions.

1. Is there something about thieving that is distinctly connected to dishonesty, in a way that other forms of immorality are not?
2. If so, what might that connection be?

For what it's worth, my guess is that there is merely a statistical connection here. A good chunk of thieving involves deception of some kind or other. But maybe there is something different to it.

Honest Promise Breaker

The first thing to do here is to think about the same kind of questions that we discussed under thieving.

1. Is there something about promise breaking that is distinctly connected to dishonesty, in a way that other forms of immorality are not?
2. If so, what might that connection be?

Here I'm even less convinced that there is a distinctive connection to dishonesty. A promises B that they'll pick up B from the airport. A forgets to set their alarm, sleeps in, and hence is not in time to pick B up from the airport. They've broken a promise and done something wrong. Have they acted dishonestly *at all*?

There is something dishonest about making a promise that one never intends to keep. And such a promise I guess typically ends up getting broken. But it's not the breaking of the promise that seems to be the dishonesty in that case; it's the flawed promise making. I'm not sure that there is any connection at all. But I'd be interested in knowing if others thought differently.

Honesty and Forthrightness

I think as I use the terms 'honest' and 'dishonest', there is an asymmetry between them concerning forthrightness. Consider the following situation.

- It is common ground that A wants to know the answer to some question Q?
- B knows the answer, and knows that the answer reflects somewhat badly on B.
- There is no reason for anyone to suspect that B knows the answer; B not saying anything about Q? isn't any kind of signal that B doesn't know the answer, because no one thinks they know.
- B is not asked whether they know, so they don't have to say that they don't know.

In such a case I think it would be honest of B to reveal the answer. I might even say it was *very honest*. But I wouldn't think it was dishonest to not reveal it. To expressly say one didn't know, or even to give people evidence that one didn't know, might be dishonest. But simply not confessing to a misdeed isn't, *I think*, a kind of dishonesty.

- Do you agree with those claims, i.e., that confessions to minor misdeeds can amount to (very) honest behaviour, but non-confessions are not in themselves dishonest?
- If so, what could explain the asymmetry here?
- Also if so, under what circumstances can mere silence amount to dishonesty? Are there any such circumstances?

Honesty and Immorality

First, a big picture question.

- Is all immorality dishonest?

I think the answer is surely no. A dislikes B. So A walks up to B one day, and punches him hard in the nose. This is bad, you shouldn't punch people. But it is hard to say that it is dishonest behaviour. So not all immoral behaviour is dishonest.

But note now that we think an argument Miller makes (twice) has to fail. Here's how he explains how thieving and promise breaking are acts of dishonesty: they amount to misrepresenting the normative features of the world. The thief represents themselves as having the right to steal the goods. And that's wrong, so it's a misrepresentation. But A misrepresents themselves as having the right to punch B in the nose. And that's wrong, so it's a misrepresentation. So this kind of implicit normative misrepresentation can't be enough to amount to dishonesty.

This is why I'm a little sceptical that Miller's story about why thieving and promise breaking are dishonest. It looks like an argument that overgenerates to 'show' that assault is dishonest.

Is Honesty Always Good?

I don't have a lot to add to what Miller says on this point, save to note an interesting historical point. The main example that gets used here is, as you might guess if you've taken any philosophy classes at all, the murderer at the door. And we get the following questions.

1. Is it good to lie to the murderer at the door?
2. Is it honest to lie to the murderer at the door?

If you think, as I do, that the answers to these questions are 'yes' and 'no', then it follows that it isn't always good to be honest. That's easy enough. But I wanted to mention a couple of things about the example.

For one thing, although we normally credit the example to Kant, it's really due to Benjamin Constant, and in particular it's from chapter 8 of his Des réactions politiques. Here is one key passage, with an automated translation._

> Le principe moral, par exemple, que dire la vérité est un devoir, s'il était pris d'une manière absolue et isolée, rendrait toute société impossible. Nous en avons la preuve dans les conséquences très directes qu'a tirées de ce principe un philosophe allemand, qui va jusqu'à prétendre qu'envers des assassins qui vous demanderaient si votre ami qu'ils poursuivent n'est pas réfugié dans votre maison, le mensonge serait un crime.
>
> The moral principle, for example, that telling the truth is a duty, if taken in an absolute and isolated way, would make any society impossible. We have proof of this in the very direct consequences of this principle by a German philosopher, who goes so far as to claim that against murderers who would ask you if your friend they are pursuing is not a refugee in your house, lying would be a crime.

Now here are two interesting things about this. Constant's book is published (and I guess written) in 1796. Kant's reply, which is only 2 pages long and starts with a long quote from Constant, comes out the following year. And Constant isn't really interested in lying. What he's interested in is how to deal with political injustice. Which is a very interesting question for a leading French intellectual to be writing about in the immediate aftermath of the Reign of Terror.

Philosophers sometimes kind of ignore the dates on things they are writing about. Oh yes, Kant wrote these things in the 1780s and these things in the 1790s and we have to know the order of these because they reflect how Kant's view developed over time. That's true, but it's also maybe relevant that during that

time the world got turned upside down and maybe that should be taken seriously in thinking about why those developments may have taken place.

There are a couple of things that could follow from this. Geneviève Rousselière has argued, in On Political Responsibility in Post-Revolutionary TImes: Kant and Constant's Debate on Lying that the whole debate should be read as an allegory for their differing views about political philosophy and the Revolution, and it isn't really about lying at all. I don't know if that's right, but it is more plausible than the other interpretations I've seen of Kant's really bizarre little paper.

But the other thing to remember is that the murderer at the door example isn't really the outlandish example that it's sometimes taken to be. For a few years there, people turning up at a door and asking for where a person who they are looking to kill might be was a kind of everyday occurrence. Now sure they weren't looking to kill them on the spot, merely drag them in front of the Revolutionary Tribunal. And the Tribunal actually acquitted a few more people than you might expect from the standard popular history. But still, describing the agents of the Revolution as murderers at the door wouldn't be the most outlandish description ever. From this perspective, the Constant-Kant debate is a modernised version of the debate in *Crito*, which I think makes it a bit more interesting than the way it is usually presented.

Constant himself is an interesting figure. He's probably less famous these days than his long-time partner, Germaine de Staël, and that's probably fair. But he's still an interesting representation of a kind of liberalism that's much more popular now than it was at the time. And to give a sense of how widely known he was, he gets name-checked near the start of Marx's *18th Brumaire* with the expectation that the reader will know precisely who he's talking about. But some of his significant works, including the one that Kant engages with in the murderer at the door paper, aren't even translated into English. Automated translations can help a bit, but probably human ones would be better. I was fascinated, for example, by this claim, at the end of a long discussion about the right way to understand the duty to tell the truth.

> C'est une idée peut-être neuve, mais qui me paraît infiniment importante, que tout principe renferme, soit en lui-même, soit dans son rapport avec un autre principe, son moyen d'application.
>
> It is perhaps a new idea, but it seems infinitely important to me, that every principle contains, either in itself or in its relationship with another principle, its means of application.

That's a really fascinating of some ideas we find playing a central role in Wittgenstein's *Investigations*. Anyway, this has been a long-enough digression, even if we did end up back somewhere from earlier in the course. Let's get back to honesty.

Instrumental Honesty

This is something we could ask about all virtues, but let's ask it about honesty. If someone consistently manifests all the signs of a virtue, but they mostly do this for instrumental reasons, do they really have the virtue? Let's think about two cases.

- A shopkeeper believes that a reputation for honesty is crucial for the success of his shop, and he believes that he isn't skillful enough to fake such a reputation. The only way he'll be seen to be honest is if he acts honestly. So that's how he acts, consistently.

- A theist believes that a reputation for honesty is crucial for getting to heaven, and she believes that she isn't skillful enough to fake such a reputation. (She doesn't really think anyone could fake out God.) The only way she'll be seen to be honest by God is if she acts honestly. So that's how she acts, consistently.

Again, a few questions.

1. Is there any difference in virtue between these two people?
2. If not (as I think), are they both honest, or both dishonest?
3. If there is a difference, what makes the cases different?

I used to think that the two cases were alike and they were both not really cases of honesty. But as I've got older and mellower, I've started to think that consistently acting out a virtue might just be enough to have it. These both seem like honest people to me. They might have gotten to honesty in a slightly sub-optimal way, but we all get to virtues in weird and roundabout ways.

But what do you think? Does honesty require having a really deep commitment to honesty? Or is doing it for the rewards, either in this life or the next, good enough?

Miller doesn't discuss this, but we could ask the same about the reverse case. Imagine someone who really wants to get ahead in a particular business. And they want this because it will mean they have more money that they can use to help raise their children, donate to charity, etc. But their manager really values dishonesty - he thinks that it's crucial to succeed in this business. So the person has to develop a reputation for dishonesty. And they don't think they can fake it - they have to actually act dishonestly. So that's what they do - feeling bad about it because they value honesty; they just value money for their children and the needy a bit more. Is this kind of instrumental dishonesty enough to make them have the vice of dishonesty? If so, is this an asymmetry between honesty and dishonesty?

Honest Mistakes

Miller leaves it as an open question whether honesty is a matter of distorting the facts as they are, or the facts as they appear. But this seems like an easy question to me. We have the concept of an honest mistake. That surely means a kind of case where someone says something false - the mistake part - but is being honest because they are saying what they believe. That is, you can't be dishonest while being completely sincere; at worst you're making an honest mistake.

I'm not normally this categorical about disputed philosophical questions, but I really don't see what the rival view could be here.

Threshold Concepts

Miller's notion of a threshold concept was, I thought, a bit confusing. And I really wasn't clear on what he meant by saying that both honesty and dishonesty are threshold concepts. This is in part because he seemed to run together two notions here. Both of them are interesting, but they aren't the same notion.

The first notion concerns the mediocre. He wants to say that they are (determinately) neither honest nor dishonest. So the person who occasionally lies to get out of a jam, every so often engages in petty theft

(office supplies and the like) or makes a promise they don't have any intent of keeping, but really does all these things less often than the median human, is neither honest nor dishonest. The point here is to deny that honest and dishonest merge into each other. Just like you might think (I guess I think) that there are shades of purple that are determinately neither red nor blue, there are dispositions vis a vis honesty that are neither on honest nor dishonest. I'm not 100% sure I agree, but it sounds plausible enough.

The second notion concerns the multi-dimensionality of the notion of honesty. Miller thinks it has these five components that he lists. I don't really think they are all part of honesty, but I do think forthrightness is part of honesty, and generally I agree that it's multidimensional. Here the notion of a threshold concept seems to be that if you're doing really badly on any dimension, that defeats calling you honest. And I guess I agree with that. If I thought non-deceptive theft was a kind of dishonesty, I would say that a frequent thief is not honest, no matter how few lies they tell etc.

But here's where the problems arise. Dishonesty doesn't have anything like the same logical structure. Someone who is really bad on one dimension of honesty/dishonesty is dishonest, even if they are really good on the other dimensions. If *being a threshold concept* requires that you score highly enough on every dimension in order to satisfy the concept, then dishonesty is not a threshold concept. And it's misleading to talk about this being a case of symmetry between honesty and dishonesty. One of them, the positive one, requires a high enough score on every dimension. The other one can be satisfied by being dishonest enough in just one way.

To some extent this is terminological. But I think it's important to note that the virtue and the vice really don't have the same logical structure here, and talking as if they do just obscures things.

Psych Problem 1: Fallacy of Composition

Much of the paper is given over to empirical evidence that people are not in fact honest. And much of that was interesting, but occasionally not as strong a support for Miller's pessimistic conclusions as he presented. The reason for this is independently interesting enough to go over.

The fallacy of composition is when one attributes to the group things that are really just true of a small number of the parts. The example that I see most commonly concerns voting behaviour. It's common to hear people say that such and such community changed their mind when the majority swung between one political group and another. Now in a democracy this change in majority might be really significant practically. But talking about it as a change in mind of the group is potentially misleading - it might have been 10% or less of the group that actually changed their mind for the majority to change.

Some of Miller's examples seemed to be like that. You could get a reported average of 6.6 heads in the coin flipping experiment by (a) a small bias, either by luck or design, in the coins (and note that they did get 5.2 not 5 in a control group), plus (b) a small-to-moderate sized group who went overboard with dishonesty. You don't need particularly widespread dishonesty to get the results.

Now to be fair, some of the tests did seem to require (or directly test for) dishonesty on a person-by-person basis. But this kind of inference from properties of salient individuals to properties of the group is a bit too common in use of empirical data, especially by non-psychologists. And you should always watch out for it.

Psych Problem 2: Replication

One big problem with using psych papers in philosophy, especially papers from before 2012 or so, is that the results of individual experiments don't always replicate. Just why this is so is a big and controversial question, but it's something you have to always be careful for.

Now Miller cites so many papers, from so many different authors, that it is unlikely that all of them will have replication problems. (Even there, you'd want to know what the connections are between the authors. If they all went to the one grad school, there could be systematic problems to watch out for.) And I suspect many of these would do fine, though it would always be nice to know.

But note that one of them, which did have a bit of uptake outside of academia, totally failed to replicate. This was the experiment that claimed people were more honest if you put the "Sign here to promise that you're honest" at the start rather than the end. And it turns out, that doesn't really work. Here's the abstract for Signing at the beginning versus at the end does not decrease dishonesty

> In 2012, five of the current authors published a paper in PNAS showing that people are more honest when they are asked to sign a veracity statement at the beginning instead of at the end of a tax or insurance audit form. In a recent investigation, across five related experiments we failed to find an effect of signing at the beginning on dishonesty. Following up on these studies, we conducted one preregistered, high-powered direct replication of experiment 1 of the PNAS paper, in which we failed to replicate the original result. The current paper updates the scientific record by showing that signing at the beginning is unlikely to be a simple solution for increasing honest reporting.

Some of the non-replications, I suspect, are due to dishonesty in the original papers. But they don't normally get critiqued by the same authors. (Though there are some prominent cases where people accuse their past co-authors of dishonesty.) Sometimes there are honest mistakes, or bad luck. And if I had to guess, I'd say that's what happened here.

I'm just noting this because the 2012 paper is one of the ones that Miller cites. It's good to bring empirical data to the question of how people actually are, but you've got to be careful about how good the data are.