

Class 17 - Intentionality

Knobe, "The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology"

I. Folk Psychology and Theory Theory

The Knobe paper describes what has come to be known as the [Knobe effect](#), though he does not call it that.

The Knobe effect is that our beliefs about the moral status of an action can influence our beliefs about whether that action was performed intentionally.

It is not absolutely clear what it means to perform an action intentionally.

There are clear cases of intentional action.

I am intentionally typing these notes, right now.

Other cases are less clear.

Typically, philosophers link intentional action with foresight and trying: I do not act intentionally when I do not foresee the consequences of my action, or if I try to avoid the action, or its consequences.

There are conflicting armchair analyses of the concept.

Knobe sought to examine how the folk use the term.

He found a discrepancy between standard philosophical analyses, which presume that whether an action is done intentionally or not is independent of the moral status of that action.

Thus, the Knobe effect was surprising.

There are two ways to resolve the tension between the standard philosophical analyses and Knobe's evidence.

First, people's beliefs about intentional actions may fail to reflect the facts about intentional actions.

People's beliefs can fail to reflect facts about intentional actions if they get them wrong, or if there are no facts.

Second, whether an action is intentional or not may depend on its moral character.

Knobe places the importance of his work in the context of the question of whether folk psychology is actually a scientific theory.

This question has various implications.

If we take folk psychology to be a theory, then we are quickly led to wonder whether it is a true (or largely true) theory.

The odd reasoning patterns we have examined this term seem particularly worrisome for the claim that folk psychology is a scientific theory.

If it is, it looks like a bad scientific theory.

Another implication of the question of whether folk psychology is a scientific theory involves the status of folk-psychological explanations of our behavior.

Folk psychology conflicts, in places, with neuroscientific explanations of behavior.

Folk-psychological explanations of mental illness yield to neurochemical analyses.

Folk explanations of dreaming and consciousness seem likely to be replaced by the analyses of a mature neuroscience.

If folk psychology is a theory, then we might well have to give it up in favor of more scientifically sophisticated theories of mind.

This is the route that eliminative materialists take in arguing, for example, that there are no beliefs, and no such thing as conscious experience, or qualia.

Just as we gave up antiquated scientific theories, like Aristotelian physics or Ptolemaic cosmology, we should give up our language of belief, intentions, and conscious awareness if we find that they do not correspond to specific neurological states.

Further, it seems implausible that there is any neural state that can serve as the meaning or reduction of my belief that, say, chocolate pudding is tasty.

Eliminativists like Patricia Churchland defend what they call theory-theory, the claim that folk psychology, like other antiquated scientific theories, is just a bad theory which should be abandoned.

In order to determine whether folk psychology is a theory, Knobe proposes two characteristics of theories: prediction and explanation of behavior.

If folk psychology is a theory, then it will be used for prediction and explanation.

If it is not used for prediction and explanation, then it is not a theory.

If it is used for purposes other than prediction and explanation, then folk psychology is not a scientific theory.

Knobe thus examines the folk-psychological concept of intentional action by asking whether intentions are used for prediction or explanation.

As he points out, intentionality is only one folk-psychological concept.

It is an open question, Knobe claims, whether the conclusions about intentionality apply to other folk-psychological concepts.

It is clear, though, that Knobe thinks that his conclusions are more widely applicable than just to this one concept.

If intentions are not a scientific concept, then we would have a good reason to believe that other folk-psychological concepts are not scientific, either, and that the role of folk psychology is not to serve as a theory used to predict and explain.

Eliminativists would not be able to defend replacing folk psychology with neuroscience.

The ramifications for Knobe's study, therefore, are deep.

II. The Knobe Effect

The Knobe effect is found, simply, in his CEO case.

Subjects were presented with a story in which a CEO approves a project to make money for his company, understanding that it will either help or harm the environment.

There were two versions of the story; in one version the project will help the environment, in the other it will harm the environment.

In both cases, the CEO says that he only cares about making money, and not whether he harms or helps the environment.

In the harm case, people attribute intention to the CEO.

In the help case, people withhold attributions of intention to the CEO.

Similar results arise when the same questions are put into the context of a lieutenant helping or harming his troops.

In [that case](#), a lieutenant orders a sergeant to take his troops either into or out of the line of enemy fire, in order to take control of a hill.

Again, in the case where the troops are harmed, people ascribe intention to the lieutenant.

In the case where the troops are helped, people do not.

The Knobe effect is surprising precisely because it is ordinarily assumed that whether we intend to perform an action is independent of the moral worth of that action.

First, we settle intentions.

Then, we evaluate actions, perhaps based on intentions.

For example, in Kantian deontological approaches to morality, we must determine our intentions in order to determine whether we can in accordance with or in violation of the categorical imperative.

Consequentialist theories of morality, like Mill's utilitarianism, are characterized by their insistence on the irrelevance of one's intentions to the morality of one's actions.

Both deontological and consequentialist theories agree that one's intentions do not depend on the morality of an act.

So, the Knobe effect is puzzling.

III. Effects and Side Effects

One might think that the question of whether one intentionally or unintentionally harms or helps the environment or one's troops is determined by analyzing (in the armchair) the concept of intentionality. Unfortunately, different analyses lead to different conclusions.

On some interpretations of the concept, the CEO's actions were intentional; on others they were unintentional.

The case in question is tricky, and leads to disparate conclusions, because it regards a side-effect.

Side effects are consequences that arise from intentionally performing an action with a different purpose.

The CEO clearly intends to increase profits.

In both the harm and help versions, the effects on the environment are side effects.

The question is whether side effects are intentional.

Some philosophers say that side effects are intentional as long as they are foreseen, or foreseeable.

Others argue that they are unintentional as long as they are not what the agent intends, or primarily intends.

But, cases can be adduced for and against either of these analyses.

I will not pursue the interesting literature on action theory, here.

Similar conflicts in our intuitions about intention arise from other kinds of complications.

In cases in which an agent lacks the appropriate skill to achieve his/her intended consequence, people's intuitions appear liable to influence by moral considerations as well.

People fail to ascribe an intention to hit the bull's-eye to Jake, in Knobe's rifle-contest case.

Jake is a lousy shot, so it is only out of luck that he hits the bull's-eye.

But, people do ascribe an intention in similar cases to Jake (in the scenario in which he kills his rich aunt) and Klaus (who sabotages his regiment's immoral mission), even when he has the same lack of skill.

In those latter two cases, like the CEO case, our moral approbation or disapprobation affects our ascriptions of intention

In the rifle-contest case, which appears morally neutral, we do not ascribe intention.

III. Intentions, Science, and Evaluation

Knobe explains the discrepancies among the various cases by dismissing the hypothesis that folk psychology is a scientific theory.

Instead, he argues, it is a theory used for attributing praise and blame.

If ascriptions of intentions were used for prediction and explanation, subjects would respond in similar ways to the harm case and the help case, and in similar ways to all three cases of the lousy shot.

But, the cases differ radically.

In the CEO case, people are more inclined to say that an agent brought about a side effect intentionally when they regard that side effect as bad than when they regard it as good.

There is an analogous asymmetry in people's judgments about praise and blame.

People are more inclined to ascribe praise or blame for a side effect when they regard it as bad than when they regard it as good.

Thus, the Knobe effect shows that people's attributions of intention track their attributions of praise and blame.

The concept of intentionality may still be useful in prediction and explanation.

But, it appears to be made for different purposes.

One might be tempted, looking at the CEO case, to say that we ascribe intentions to agents when they act in ways that harm, but not when they act in ways that help.

Knobe's data support a more subtle hypothesis.

For, the Klaus case, in which people ascribe intentions to Klaus when he sabotages an immoral mission, shows that ascriptions of intention may be made in cases when good consequences result.

The key to the difference is that people want to praise Klaus, but they do not want to praise the CEO.

People's intentional action intuitions tend to track the psychological features that are most relevant to praise and blame judgments. But...different psychological features will be relevant depending on whether the behavior itself is good or bad. That is to say, we use different psychological features when we are (a) trying to determine whether or not an agent deserves blame for her bad behaviors from the ones we use when we are (b) trying to determine whether or not an agent deserves praise for her good behaviors (144).

It is clear that people's sensitivity to the moral characteristics of an action affect their attributions of intentionality.

Thus, the concept of intentionality, and thus the concepts of folk psychology generally, seem unfit for prediction and explanation.

That is, the Knobe effect supports his claim that folk psychology is not a theory.

Knobe makes the point better [elsewhere](#) (see §10):

Suppose that we are observing a hunter who is trying to shoot a deer. The hunter is extremely unskilled, but - through sheer luck - he manages to hit the deer directly in the heart. Our task now is to determine whether or not the hunter killed the deer intentionally. According to the hypothesis defended here, our judgment will be influenced by the degree to which we believe that the hunter deserves reduced praise or blame because of his lack of skill. Thus, if some of us see the killing of deer as morally wrong and others see it as an achievement, the final result will probably be that some of us see the behavior as intentional and others do not.

It is hard to see any way in which this difference in our judgments could help us to pursue “scientific” tasks like predicting future behavior. On the contrary, it seems that we should be able to reach perfect agreement about how our observation of the hunter should affect predictions of his future behavior even if we disagree radically about the moral status of the event we have observed (Knobe, “Intentional action in folk psychology: an experimental investigation,” 322).

Knobe’s claim is that our predictions of the hunter’s future behavior will differ depending on whether I ascribe intentions to him.

Is he correct?

IV. A Query

One further question I have is whether attributions of praise and blame themselves could be tools for prediction and explanation.

Then, perhaps, we could see the asymmetric attributions of intention still as part of a theory of folk psychology, subordinate to more explanatory or predictive evaluative hypotheses.

This claim would take some work to flesh out.