MORALITY OR MODALITY?
WHAT DOES THE ATTRIBUTION OF INTENTIONALITY DEPEND ON?

BENCE NANAY

It has been argued that the attribution of intentional actions is sensitive to our moral judgment. I suggest an alternative, where the attribution of intentional actions depends on modal (and not moral) considerations. We judge a foreseen side-effect of an agent's intentionally performed action to be intentional if the following modal claim is true: if she had not ignored considerations about the foreseen side-effect, her action might have been different (other things being equal). I go through the most important examples of the asymmetry in the attribution of intentionality and point out that the modal account can cover all the problematic cases, whereas the moral account can't.

I. Morality and the attribution of intentional actions

It has been argued that the attribution of intentional actions is sensitive to our moral judgment (Knobe 2003, 2004, 2006, 2007, see also Mele 2003, Phelan & Sarkissian 2008). I will examine these arguments and suggest an alternative explanation for the experiments they are based on.

Joshua Knobe conducted the following experiment (Knobe 2003) to support this claim. Subjects were given two vignettes that differed only in one small detail and this difference influenced their attribution of intentionality. The first vignette was the following:

> The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.'

> The chairman of the board answered, 'I don't care at all about harming the environment. I just want to make as much profit as I can. Let's start the new program.'
>
> They started the new program. Sure enough, the environment was harmed.

The second vignette differed in only one word: 'harming' was replaced by 'helping':

> The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, but it will also help the environment.'
>
> The chairman of the board answered, 'I don't care at all about helping the environment. I just want to make as much profit as I can. Let's start the new program.'
>
> They started the new program. Sure enough, the environment was helped.

The surprising finding was that the overwhelming majority (77%) of subjects said in response to the first vignette that the chairman harmed the environment intentionally, whereas the overwhelming majority (70%) of subjects said that the chairman did not help the environment intentionally. As the difference between the first and the second scenario was only in terms of the moral value of harming versus helping the environment, Knobe concludes, this experiment clearly shows that our moral judgment influences our attribution of intentional actions. Knobe and his collaborators conducted a number of similar experiments that yielded similar results (see Knobe 2006 for a summary).

But does this experiment really show that our moral judgment influences our attribution of intentional actions? I will argue that there is a simpler way of explaining these findings, without talking about moral judgment or morality at all. In short, I will argue that the difference between the two vignettes above is not a difference in the moral value of a certain action.

II. A modal account of the attribution of intentionality

One striking feature of the experiments Knobe and his collaborators conducted is that they all share the same structure. To put it very simply, in one scenario, the agent has two reasons for performing a certain action and ignores one of these. In the other, the agent has a reason for and a reason against performing an action and ignores the reason against. Thus, in the experiment I quoted above, we have the following two scenarios:

> (a) In the harm case, the chairman has a reason (R_1) for introducing the plan (to increase profit) and a reason (R_2) against (to avoid harming the environment).
>
> (b) In the help case, in contrast, the chairman has two different reasons to introduce the plan: he had a reason to increase the company's profit (R_1) and he also had a reason to help the environment (R_3).

In short, the difference between (a) and (b) is that in (a) the chairman has R_1 for and R_2 against introducing the plan, whereas in (b) he has R_1 and R_3 both in favor of performing

this action. Importantly, the chairman chooses to *ignore* the environmental considerations: R_2 and R_3, respectively. This leaves R_1 in both scenarios, which is a reason for introducing the plan. There is no difference between (a) and (b) in the actual reason the chairman is acting on.

But there is a modal difference between (a) and (b): there is a difference in what would happen if the chairman *did not ignore* R_2 and R_3, respectively. Contrast the original scenarios (a) and (b) with another pair of cases where the chairman chooses *not* to ignore the environmental considerations.

> (a*) The chairman chooses not to ignore R_2 (a reason against introducing the plan). Then his action would, or at least it could, be different, as now he has a reason for (R_1) and a reason against (R_2) introducing the plan.

> (b*) The chairman chooses not to ignore R_3 (a reason for introducing the plan). His action *would still be the same*, as now he has two reasons (R_1 and R_3) in favor of introducing the plan)

So an important difference between case (a) and case (b) is a modal one: the outcome would be different if the chairman didn't ignore the environmental considerations. In (b), ignoring that the plan helps the environment would make no difference, as there are two independent reasons in favor of introducing the plan: the chairman's action in (b) and (b*) will be the same. In (a), on the other hand, ignoring that the plan harms the environment would make (or at least it could make) a difference: the chairman's action in (a) and (a*) will be (or at least can be) different.

Thus, what this experiment shows is that in (b) introducing the new scheme does not depend counterfactually on ignoring the environmental considerations, whereas in (a), there is counterfactual dependence between ignoring the environmental considerations and introducing the new scheme. This counterfactual dependence in (a) is not very strong, as not ignoring will not guarantee that the chairman's action will be different, but it is an instance of counterfactual dependence nonetheless. In (b), we have no counterfactual dependence, weak or strong.

So far, I gave an account of *how* the attribution of intentionality to someone else's action works. Now I want to give a sketchy outline of *why* it works this way. The alternatives of the account I proposed have not given any justification for why the attribution of intentionality to someone else's action works the way it does. They are exclusively interested in answering the how question. I, however, would also like to at least sketch a possible answer to the why question. The sketchiness of this answer needs to be emphasized: a full explanation would be a subject for another paper.

It is important to note that the chairman ignores the environmental considerations in both (a) and (b) *intentionally*: he knows perfectly well what it is that he is ignoring and he chooses to ignore it anyway. Thus, in (a) the chairman's action depends counterfactually on his intentional mental action (of ignoring environmental considerations), whereas in (b) it does not.

Thus, in (a), the chairman's harming the environment depends counterfactually on an intentional mental action of his with the content 'harming the environment'. The chairman's Q-ing depends counterfactually on an intentional mental action with the content 'to Q'. It is very tempting to conclude on the basis of this claim that the chairman harmed the environment intentionally. Here is why.

First, although counterfactual dependence is, of course, neither necessary nor sufficient for causation, it is really easy to confuse the two. Thus, if we know that the chairman's Q-ing depends counterfactually on his intentional mental action with the content 'to Q', it is easy to conclude that the fact that the chairman's intentional mental action was causally responsible for his Q-ing.

Second, according to a widespread conception of intentional action, popular both among philosophers and among non-philosophers, an action Q is intentional if it is caused by an intention to Q. The chairman's harming the environment was *not* caused by his intention to do so, but it was caused by something very similar: an intentional mental action with the content 'to Q'. Thus, if we read vignette (a), it is easy to slide from counterfactual dependence to causal responsibility and it is also easy to slide from being caused by an intentional action with the content 'to Q' to being caused by the intention to Q, even for a philosopher, let alone for the non-philosophers who were the subjects of Knobe's experiments.

Thus, in (a), the counterfactual dependence between the chairman's harming the environment and his ignoring environmental considerations is responsible for why we are tempted to put down his action as intentional. In (b), there is no such counterfactual dependence; hence, there is no temptation to interpret his action as intentional either. The difference between (a) and (b) is not a moral, but a modal one.

III. Reason attributions

The modal explanatory scheme I outlined in the last section accounts for Knobe's results without appealing to moral judgments. But it does appeal to reason attributions: we attribute

reasons to the chairman and whether or not we take him to perform an action intentionally depends on the way these reasons weigh against one another.

The claim that the attribution of intentionality depends on the attribution of reasons is by no means new. Gilbert Harman famously wrote that "one can do something intentionally even though one does not intend to do it, if one does it in the face of what ought to be a reason not to do it and, either one tries to do it, or one does it as a foreseen consequence of something else that one intends to do" (Harman 1979, p. 434).

But the reliance on reason attribution does raise a couple of important questions that need to be clarified before proceeding. First of all, what does reason attribution really mean?

Harman talks about the way in which the intentionality of one's action depends on one's reasons to do such and such. But what we are trying to explain here is not the intentionality of one's action, but the attribution of intentionality to someone else's action: what makes a person (A) judge that another person (B) acts intentionally. And my claim is that this depends on A's attribution of reasons to B. But it is not clear what A's attribution of a reason to B would mean.

There are two importantly different kinds of reason attribution:

(i) A can attribute B a reason to do something if A thinks that this is what B ought to do (although B may not think so).[1]

[1] It is worth noting that (i) itself covers two different types of reason attributions: A may attribute a reason to B because A believes that this is what would be consistent with B's beliefs and desires (although B does not recognize this). Or A may attribute a reason to B

(ii) A can attribute B a reason to do something if A thinks that this is what B thinks B ought to do (although A may not think so).

Examples: I am an atheist, but my friend, Bill, is very religious and he goes to church every week. Bill just cheated on his wife. I may attribute a reason to him to go and confess his sin: I may think that he has a reason to go to church, that is to say, I may think that this is what Bill thinks he ought to do. I definitely do not think that this is what he ought to do: I'm an atheist, after all: confessing sins in church is a waste of time. This is an example for reason attribution of type (ii).

Contrast this case with the following. Bill, again, is very religious and I am, again, an atheist. Now I just cheated on my wife. He may attribute a reason to me to go and confess my sin, as he may think that this is the only way I can avoid going to hell. But he knows that I don't think so. So Bill attributes a reason to me to confess my sin. What this means is that he thinks that this is what I ought to do, but he knows that this is not what I think I ought to do. This would be an example for reason attribution of type (i).

As I think that both of these two types of reason attribution are equally genuine, I would like to allow for both in my account.[2] The difference between the two will play an important role in the next section.

---

independently of B's beliefs and desires. I take both of these two kinds of reason attributions to be genuine.

[2] This is a major difference between my account and the one Frank Hindriks gives. According to him, "An agent S As intentionally if S intends to B, As by B-ing, expects to A by B-ing, and […] Ss Bs in spite of the fact that she believes her expected A-ing constitutes a

The second question that needs to be addressed about reason attribution is what our reason attributions depend on. If reason attributions depend on our moral judgment, then my explanatory scheme is not really different (although it may be more detailed) than Knobe's.

Reason attributions are in fact sometimes sensitive to our moral judgments. Whether I think that an agent has a reason to do something sometimes depends on what *I* judge to be moral or immoral, as in the confession example for reason attribution of type (i) above. Sometimes it depends on whether I think *she* judges it to be moral or immoral, as in the confession example for reason attribution of type (ii) above. But, importantly, sometimes it does not depend on any moral considerations at all. I may think that Bill has a reason to order pancakes rather than toasts if I know his food preferences (or food allergies) well enough, but this reason attribution does not depend on any moral judgment. As reason attribution does not always depend on our moral judgments, if the argument I gave in the last section is correct, then the attribution of intentional actions does not depend on our moral judgments either.

IV. Morality vs. modality: four test cases

---

normative reason against her B-ing" (Hindriks forthcoming, p. 10). Hindriks's Harman-inspired analysis would only allow for reason attribution of type (ii), which makes his solution to the Knobe problem vulnerable to counterexamples of the kind I review in Section IV. (c).

We have two ways of explaining Knobe's original experiments: one brings in moral judgments, the other one does not. So far my account does not fare better than Knobe's. In what follows, I will argue that (i) some of Knobe's own experiments can be explained in a more straightforward and less *ad hoc* manner if we use my account and (ii) there are some experiments that can only be explained with the help of my account.

Of course there are other attempts to explain the asymmetry of the attribution of intentionality besides Knobe's and my own. One important example is Edouard Machery's 'trade off' hypothesis. I will review some of the experiments that are supposed to support and falsify Machery's account and examine whether Knobe's and my explanatory scheme would fit them.

IV. (a) The first problem case: Racial identification laws

The first experiment I want to appeal to is one of Knobe's own (Knobe 2007). The first of the two scenarios that are being compared here is the following:

> In Nazi Germany, there was a law called the 'racial identification law.' The purpose of the law was to help identify people of certain races so that they could be rounded up and sent to concentration camps.
>
> Shortly after this law was passed, the CEO of a small corporation decided to make certain organizational changes.
>
> The Vice-President of the corporation said: "By making those changes, you'll definitely be increasing our profits. But you'll also be violating the requirements of the racial identification law."

> The CEO said: "Look, I know that I'll be violating the requirements of the law, but I don't care one bit about that. All I care about is making as much profit as I can. Let's make those organizational changes!"
>
> As soon as the CEO gave this order, the corporation began making the organizational changes.

And here is the second scenario:

> In Nazi Germany, there was a law called the 'racial identification law.' The purpose of the law was to help identify people of certain races so that they could be rounded up and sent to concentration camps.
>
> Shortly after this law was passed, the CEO of a small corporation decided to make certain organizational changes.
>
> The Vice-President of the corporation said: "By making those changes, you'll definitely be increasing our profits. But you'll also be fulfilling the requirements of the racial identification law."
>
> The CEO said: "Look, I know that I'll be fulfilling the requirements of the law, but I don't care one bit about that. All I care about is making as much profit as I can. Let's make those organizational changes!"
>
> As soon as the CEO gave this order, the corporation began making the organizational changes.

The vast majority (81%) of the subjects said that the CEO intentionally violated the requirements of the racial identification law (in the first case) and the vast majority (70%) of

the subjects said that the CEO did not intentionally fulfill the requirements of the racial identification law (in the second case). This example presents some difficulties to Knobe's account: the attribution of intentional actions does not depend on what we, the evaluators, judge to be moral, as it seems reasonable to assume that the subjects who were confronted with these scenarios take the racial identification law to be something morally despicable. Knobe fixes his account by suggesting that we make a non-conscious moral judgment that goes against our all things considered conscious moral judgment. And it is the non-conscious moral judgment that our attribution of intentional actions depend on (Knobe 2007).

But my account can explain these results without the slightly *ad hoc* postulation of 'non-conscious moral judgments'. If we accept my account, then subjects attribute a reason to the CEO for introducing the organizational changes in both scenarios: to make more profit. And, in the case of the first vignette, they also attribute a reason to the CEO against violating the law. We can attribute this reason to the CEO and still think that the law is morally despicable: the CEO lived in terrible times and in those times people had reason to fulfill morally despicable laws, if they want to stay out of prison.

As in the case of the first vignette, the CEO has a reason for and a reason against introducing the organizational changes, ignoring the latter can make a difference with regards to the outcome of his action. And as we have seen, this makes it tempting to conclude that he violated the racial identification law intentionally. In the case of the second vignette, ignoring the reason to play along with the racial identification law would make no

difference with regards to the outcome of the CEO's action, hence, there is no temptation to conclude that the CEO fulfilled the requirements of the law intentionally.[3]

[3] It is of course an oversimplification to say that we attribute only two reasons to the CEO: one (profit-related) reason for introducing the plan and one (racial identification law-related) reason against. We presumably attribute a number of reasons to the CEO. We may attribute a reason to him for violating the racial identification law on moral grounds. And we may attribute a reason to him against violating this law on the grounds that if he did so, he would easily end up in prison. These two reasons clearly pull in different directions. If we assume that we attribute not two, but three different reasons to the CEO, we get the following picture. In the first scenario, we attribute one reason for breaking the law (making profit) and two reasons against it (taking risks and a moral reason). In the second scenario, we attribute two reasons for complying with the law (making profit and not taking the risk) and one reason against (a moral reason). So it may seem that my account implies that in the second scenario people should say that the CEO intentionally complied with the law. But in fact they do not say this, so I have some explaining to do. It is important that my account does not imply the conclusion that in the second scenario people should say that the CEO intentionally complied with the law. According to my account, judging that the CEO intentionally complied with the law depends on taking the CEO to be intentionally ignoring a reason to break the law. So in the second scenario we may attribute a reason to the CEO against complying the law (the moral reason), but we do not take the CEO to be intentionally ignoring this reason, as nothing in the text indicates so. And if we do not take the CEO to be intentionally ignoring a reason against complying with the law, we do not

Thus, my account can explain cases when we attribute reasons to agents that we wouldn't necessarily endorse. Knobe's account, in contrast, needed to be modified to cover cases like this.

IV. (b) The second problem case: Free-cup versus extra dollar

Edouard Machery argued that what explains the asymmetry of the attribution of intentionality in the original Knobe case is a 'trade off' hypothesis: when we judge an action to be intentional, we do so because we construe the bad side effect as a cost that must be assumed in order to achieve something else (Machery 2008). In the original example, hurting the environment is the cost of increasing profit. There is a trade-off between the two. But helping the environment is not the cost of increasing profit: there is no trade-off here.

Machery's explanatory scheme is similar to mine in as much as it purports to explain the Knobe effect without any appeal to morality. It is important to examine how my account would explain the examples he gives in support of his 'trade-off hypothesis'. In the next section, I will examine some examples that Machery's account cannot accommodate.

Here is Machery's example for an asymmetry of the attribution of intentionality that has nothing to do with morality (Machery 2008, p. 177). The first scenario is the following:

> Joe was feeling quite dehydrated, so he stopped by the local smoothie shop to buy the largest sized drink available. Before ordering, the cashier told him

---

take him to be intentionally complying with the law. (I am grateful to an anonymous reviewer both for raising this extra complication and also for suggesting a way to resolve it.)

> that if he bought a Mega-Sized Smoothie he would get it in a special commemorative cup. Joe replied, 'I don't care about a commemorative cup, I just want the biggest smoothie you have.' Sure enough, Joe received the Mega-Sized Smoothie in a commemorative cup.

The contrasting scenario was this:

> Joe was feeling quite dehydrated, so he stopped by the local smoothie shop to buy the largest sized drink available. Before ordering, the cashier told him that the Mega-Sized Smoothies were now one dollar more than they used to be. Joe replied, 'I don't care if I have to pay one dollar more, I just want the biggest smoothie you have.' Sure enough, Joe received the Mega-Sized Smoothie and paid one dollar more for it.

The subjects were asked whether Joe intentionally obtained the commemorative cup and whether Joe intentionally paid one dollar more, respectively. 95% of respondents said yes in the extra dollar scenario, whereas only 45% said yes in the commemorative cup case. Thus, there is a difference between the two scenarios in terms of the attribution of intentionality, but this difference cannot be due to differences with regards to the moral value of the side effects, as there is nothing moral or immoral about either side effect.

Machery argues that this asymmetry could be explained with the help of his 'trade-off hypothesis': the extra dollar is seen as the cost of, and thus, is traded off against, the delicious smoothie, whereas the commemorative cup is not.

It is important to note that my account can also explain this asymmetry. In the extra dollar case, Joe has two reasons: one for buying the smoothie, one against. He has a strong reason for: he is dehydrated. And he has a not so strong reason against: one buck out of his pocket. He decides to ignore the latter reason. If he had not done so – if he had been stingier –, he would probably not have bought the Mega-Sized Smoothie, maybe settling for a smaller drink. So ignoring his 'extra dollar' reason could have made a difference to the outcome of his actions: as we have seen in Section II, in cases where there is such counterfactual dependence, we are tempted to attribute intentionality to the action the agent chooses to ignore.

Contrast this with the commemorative cup case, where Joe has a strong reason for buying the smoothie, as before, but he does not have any reason to get the commemorative cup, or, maybe if he is bored, he has some (weak) reason for getting it. And he decides to ignore considerations of the commemorative cup. But if he did not ignore these considerations, he would still have bought the Mega-Sized Smoothie. So this act of ignoring would not have made any difference to the outcome of his actions. And in this case we will not be tempted to say that Joe obtained the commemorative cup intentionally.

Thus, my account can explain this example as well as Machery's, and, as Machery rightly points out, Knobe cannot do so.[4] My account seems to fare better than Knobe's. But

[4] My account can also explain why people with Asperger syndrome judge that the side effect is intentional in both the free cup case and in the extra dollar case (Zalla and Machery ms). It follows from my account that people with Asperger syndrome fail to attribute a reason not to pay the extra dollar. And this explanation is consistent with the general understanding of

now the worry becomes whether my account enjoys any explanatory advantage over Machery's 'trade-off hypothesis'.

IV. (c) The third problem case: The Gang Leader

Machery's 'trade-off hypothesis' may seem very similar to the account I proposed above: after all while Machery talks about trade off, I talk about the conflict between a reason for and a reason against performing an action. Thus, it is important to examine whether my account is vulnerable to the most serious objection Machery's 'trade-off hypothesis' faces.

Ron Mallon gave a case that is unlikely to be accommodated by Machery's 'trade-off hypothesis' (Mallon 2008, p. 252):

> A member of a local gang went to the leader and said, "We are thinking of trying a new tactic. It will flood the neighborhood with cheaper cocaine, increasing our profits, but it will also harm the cops since more cops will die in drug-related violence."
>
> The leader answered, "I admit that it would be good to harm the cops, but I don't really care about that. I just want to make as much profit as I can. Let's implement the new tactic."
>
> They did implement the new tactic, and sure enough, the cops were harmed since more cops died in drug-related violence.

Asperger syndrome, as the most salient difference between people with and without Asperger syndrome concerns the attribution of mental states to others.

The other vignette was the following:

> A member of a local gang went to the leader and said, "We are thinking of trying a new tactic. It will flood the neighborhood with cheaper cocaine, increasing our profits, but it will also help hard-core addicts to have more money for food and housing."
>
> The leader answered, "I admit that it would be good for people to have more money for food and housing, but I don't really care about that. I just want to make as much profit as I can. Let's implement the new tactic."
>
> They did implement the new tactic, and sure enough, hard-core addicts were helped by having more money.

The subjects were asked whether the leader of the gang intentionally harmed the cops and whether the leader of the gang intentionally helped hard-core addicts, respectively. 62% of respondents said yes in the cop-harming scenario, whereas only 28% said yes in the addict-helping case. Thus, there is a difference between the two scenarios in terms of the attribution of intentionality, but this difference cannot be due to any trade-off, as neither harming cops, nor helping addicts could be construed as the cost of achieving more profit. There is no trade off in either scenario, yet, there is a statistically significant difference in terms of the attribution of intentionality. Bad news for Machery's 'trade-off hypothesis.

My account, in contrast, can explain this asymmetry. In the harming cops case, we attribute two reasons to the gang leader: one for flooding the neighborhood with cheaper cocaine, one against. He has a strong reason for: he wants more profit. And we attribute a

reason against: he ought not to harm cops – maybe because we think harming cops is immoral, or maybe because we think it is in his interest not to harm cops, in order not to provoke more police interference. But we know that the gang leader does not think this way. This is an instance of reason attribution of type (i): we attribute this reason to the gang leader because we think this is what he ought to do, in spite of the fact that we know that this is not what he himself thinks. But as we have seen in Section III, this still counts as a reason attribution: A can attribute B a reason to do something if A thinks that this is what B ought to do (although B may not think so).

The gang leader decides to ignore this reason – which, again, he does not see as a reason against, but as a (weak) reason for, introducing the new plan. If he had not done so – if he had not ignored what we think he should have done –, maybe he would not have flooded the neighborhood with cheap cocaine. So ignoring his 'harming cops' reason could have made a difference to the outcome of his actions. And, as we have seen in Section II, in cases where there is such counterfactual dependence, we slip easily into attributing intentionality to the action the agent chooses to ignore.

Contrast this with the 'helping addicts' case, where the gang leader has a strong reason for flooding the neighborhood with cheaper cocaine, as before, but we do not attribute any reason to him against this course of action. We attribute another reason for introducing the plan: we think helping addicts is good, and, according to the vignette, he also thinks so. And he decides to ignore one of his reasons for introducing the new plan (the consideration for helping addicts). But if he did not ignore these considerations, he would still have introduced the new plan. So this act of ignoring would not have made any difference to the outcome of his actions. And in this case we will not be tempted to say that the gang leader helped hard core addicts intentionally.

Thus, my account can explain the case that casts doubt over Machery's 'trade-off hypothesis'.

IV. (d) The fourth problem case: The apple tree

But there are some further examples that my explanatory scheme can accommodate, while Knobe's is unlikely to do so. Here is a new one. The difference between the two scenarios is very similar to the one between Knobe's two original vignettes, as far as the reason attributions are concerned, but there is no moral value involved in either of them. As the outcomes are also similar to Knobe's original experiment, this experiment suggests that the attribution of intentional action does not depend on our moral judgments, conscious or non-conscious. The first vignette is the following:

> The vice-president of a company went to the chairman of the board and said, 'We are thinking of expanding our building. It will help us increase profits, but it will also mean that we have to cut down the apple tree in front of your office.'
>
> The chairman of the board answered, 'Although I spent a lot of time as a child climbing on that apple tree and I enjoyed its view from my office, I don't care at all about cutting it down. I just want to make as much profit as I can. Let's start the expansion.'
>
> They started the expansion. And the apple tree was removed.

The second vignette tells the same story but the chairman has slightly different attitude towards the apple tree:

> The vice-president of a company went to the chairman of the board and said, 'We are thinking of expanding our building. It will help us increase profits, but it will also mean that we have to cut down the apple tree in front of your office.'
>
> The chairman of the board answered, 'Although that apple tree has annoyed me ever since I moved into this office, I don't care at all about cutting it down. I just want to make as much profit as I can. Let's start the expansion.'
>
> They started the expansion. And the apple tree was removed.

The overwhelming majority (78%) of subjects said that the chairman in the first scenario got rid of the tree intentionally and the vast majority (69%) of subjects said that the chairman in the second scenario did not get rid of the tree intentionally.[5] But there is no difference with regards to the moral status of the actions the chairman performs in the two scenarios: he performs the same action. Where the two scenarios differ is that the chairman had a reason not to perform the action in the first scenario whereas he had a reason to perform it in the second. As neither of these reasons is moral in character, we have a difference with regards to the attribution of intentional action without any difference with regards to morality.

[5] 40 subjects were asked. They were all my students in the fourth year philosophy of perception course at the University of British Columbia.

Again, my account can explain this difference in modal terms. In the first scenario, ignoring the sentimental reasons against getting rid of the apple tree would (or at least could) have made a difference with regards to the chairman's action. In the second scenario, ignoring the personal reasons for getting rid of it would not make any difference. And this modal difference is what explains the difference in our attribution of intentional actions, not moral considerations, conscious or unconscious.

V. Conclusion

I have considered four problem cases in the last section: Knobe's, Machery's, Mallon's and my own. None of the existing accounts can cover all of these (Shaun Nichols and Joseph Ulatowski also noted that no existing account can cover all the experiments: Nichols & Ulatowski 2007, p. 351). Knobe has only an *ad hoc* explanation of the first and no explanation for the second and the fourth. Machery has no explanation for the third and the fourth. I aimed to point out that the modal account I proposed can explain all four of them. The conclusion then is that the attribution of intentionality has little to do with morality, but a lot to do with modality: with what would have happened if the agent had not ignored some of her reasons.[6]

References:

[6] I am grateful for Joshua Knobe's help, comments and encouragement. I am also grateful for the extremely thorough and thoughtful comments of two referees for this Journal.

Harman, Gilbert 1979. "Practical reasoning." *Review of Metaphysics* 29 (1979): 431-463.

Hindriks, Frank forthcoming "Intentional action and the praise-blame asymmetry." *Philosophical Quarterly*.

Knobe, J. 2003. "Intentional Action and Side Effects in Ordinary Language." *Analysis,* 63 (2003): 190-193.

---. 2004. "Intention, Intentional Action and Moral Considerations." *Analysis* 64 (2004): 181-187.

---. 2006. "The concept of intentional action. A case study in the uses of folk psychology." *Philosophical Studies* 130 (2006): 203-231

---. 2007. "Reason explanation in folk psychology." *Midwest Studies in Philosophy* 31 (2007): 90-107.

Machery, Edouard 2008. "The folk concept of intentional action: Philosophical and experiential issues." *Mind & Language* 23 (2008): 165-189.

Mallon, Ron 2008. "Knobe vs. Machery: Testing the trade-off hypothesis." *Mind & Language* 23 (2008): 247-255.

Mele, A. 2003. "Intentional Action: Controversies, Data, and Core Hypotheses." *Philosophical Psychology*, 16 (2003): 325-340.

Nichols, Shaun & Joseph Ulatowski 2007. "Intuitions and individual differences: The Knobe effect revisited." *Mind & Language* 22 (2007): 346-365.

Phelan, Mark & Hagop Sarkissian 2008. "The folk strike back: Or, why you didn't do it intentionally, though it was bad and you knew it." *Philosophical Studies* 138 (2008): 291-298.

Zalla, Tiziana & Edouard Machery ms. "The concept of intentional action in Asperger syndrome." Manuscript.