

RUNNING HEAD: The Knobe Effect

The Knobe Effect: A Brief Overview

Adam Feltz

Department of Philosophy

Florida State University

Tallahassee, FL 32306-1500

adf04@fsu.edu

Abstract

Joshua Knobe (2003a) has discovered that the perceived goodness or badness of side effects of actions influences people's ascriptions of intentionality to those side effects. I present the paradigmatic cases that elicit what has been called the Knobe effect and offer some explanations of the Knobe effect. I put these explanations into two broad groups. One explains the Knobe effect by referring to our concept of intentional action. The other explains the Knobe effect without referring to our concept of intentional action. I discuss some problems with these explanations and conclude with some possible avenues for future research.

Experimental philosophy is the field of philosophy that is defined by the “use of the methods of experimental psychology to probe the way people think about philosophical issues and then examine how the results of such studies bear on traditional philosophical debates” (Nadelhoffer and Nahmias, 2007, p. 123). One area where experimental philosophy has had an impact is action theory. According to Alfred Mele, “Central to the philosophy of action is a concern to understand *intentional* action” (1992, p. 199). The contribution of experimental philosophy to traditional action theory is important because it can help shed light on what the folk concept of intentional action is. As Mele correctly points out, “a philosophical analysis of intentional action that is wholly unconstrained by that [folk] concept runs the risk of having nothing more than a philosophical fiction as its subject matter” (2001, p. 27). In order to assure that philosophers' analyses are so constrained, experimental philosophy has a unique role to play in providing empirical data on which philosophers can theorize.

In this endeavor, Joshua Knobe (2003a) has discovered that the goodness or badness of side effects influences people's intentional action intuitions. In what follows, I present the paradigmatic cases that elicit what has been called the Knobe effect and offer some possible explanations why people respond as they do. I put these explanations into two broad groups. One group explains the Knobe effect by referring to our concept of intentional action. The other group explains the Knobe effect without referring to our concept of intentional action. I discuss some problems with these explanations and conclude with some avenues for future research.

The Knobe Effect

If a consequence of an action is foreseen but not intended, then that consequence is a side

effect. The “Knobe effect” is the phenomenon where people tend to judge that a bad side effect is brought about intentionally, whereas a good side effect is judged not to be brought about intentionally. The best known cases used to demonstrate the Knobe effect are Knobe's two chairman cases:

Harm

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.

Help

The vice-president of the company went to the chairman of the board and said, “We are thinking of starting a new program. It will help us increase profits, and 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.

(Knobe, 2003a, p. 191).

About 82% of the participants given harm say that the chairman brought about the bad side effect (harming the environment) intentionally, while 77% of those given help said the chairman did *not* bring about the good side effect (helping the environment) intentionally (Knobe, 2003a, p. 192). These results have been replicated across a variety of cases involving side effects (Cushman and Mele, in press; Knobe, 2003a, 2003b, 2004a, 2004b; Knobe and Mendlow, 2004;

McCann, 2005; Mele and Cushman, 2007; Nadelhoffer, 2004a, 2004c 2005, 2006a, 2006b, 2006c), cultures (Knobe and Burra, 2006), and ages (Leslie, Knobe, and Cohen, 2006).

Given that the Knobe effect is a robust phenomenon, what explains it? There are two general explanations. One employs the folk concept of intentional action, and one does not. Exploring these different explanations is the task of the next two sections.

Competency Explanations of the Knobe Effect

Single Core Concept Explanation

Knobe (2006) attributes the asymmetric results to the core folk concept of intentional action. According to Knobe, the folk concept of intentional action is sensitive to the nature of side effects. Because the folk concept is sensitive to the nature of the side effect, those differences can affect judgments about the intentionality of the side effect.

Knobe's model of intentional action ascription features two sub-processes. First, one determines if the behavior is good or bad. Second, one's concept of intentional action is engaged using the determination from the first step to issue an intentionality judgment. For example, consider judgments made about harm. One first determines the side effect to be bad. Then, one's concept is engaged and one searches for features of the bad side effect sufficient to judge it being brought about intentionally. For example, foresight might be sufficient to judge a harmful side effect is brought about intentionally. Hence, people think the chairman intentionally harmed the environment. However, a different set of considerations might be relevant when the chairman helps the environment—foresight may no longer be sufficient. What may be required in addition to, or instead of, foresight is that the chairman had the intention or desire to help the environment. Both are lacking in help. So, they judge that the chairman did not help the

environment intentionally. Thus, “moral considerations are playing a helpful role in people's underlying competence itself” (Knobe, 2006, p. 226).¹

Phelan and Sarkissian (in press) find that the moral valence of the side effect does not fully explain the Knobe effect. In a study based on scenarios used by Knobe and Mendlow (2004), participants are given scenarios describing a president of a corporation who intends to increase sales in Massachusetts and foresees, but doesn't care, that it will decrease sales in New Jersey. Consistent with harm, most people judged that the president lowered sales in New Jersey intentionally. However, when participants are asked directly about the badness of the side effect, most participants say that the side effect is not bad. Hence, it does not look like the badness of the side effect completely explains the asymmetry.

There is evidence that regretfully bringing about a bad side effect reduces judgments of intentionality. Sverdlik (2004) ran an experiment where Jones has to mow his lawn early in the morning. He also foresees that doing so will have the unintended effect of waking up his neighbors. He regrets that, but he mows the lawn anyway thereby waking up his neighbors. Participants were also given a “non-regret” version of the scenario that does not mention whether Jones regrets waking his neighbors or not. Intentionality ratings were significantly lower in the regret case than in the non-regret case even though the bad side effects are the same. This result has been replicated by Phelan and Sarkissian (in press). Hence, again, the badness of the side effect does not completely explain intentionality judgments.

Multiple Core Concepts Explanation

Nichols and Ulatowski (2007) think that Knobe's results are best explained by what they

¹ Knobe has since abandoned this view.

call “interpretative diversity.” They used a within-participants design giving participants both harm and help. Nichols and Ulatowski replicated Knobe's asymmetry, and they, like Knobe, had a substantial number of participants dissenting from the majority responses. In fact, they found that roughly a third responded “no” to both harm and help, a third responded “yes” to harm and help, and a third responded “yes” to harm and “no” to help. Participants were asked to explain their answers. These explanations fell into two categories: (a) the chairman lacked a desire to bring about the side effect, or (b) the chairman foresaw that the side effect would be brought about. Those who judged the chairman brought about the side effect intentionally justified their answers with (b). Those who did not think the chairman brought about the side effect intentionally justified their answers with (a). Hence, Nichols and Ulatowski conclude that “considerations of outcome may influence which interpretation the term is given” and those interpretations are reflective of two separate concepts of intentional action—a knowledge based concept and a motive based concept (2007, p. 361).²

One problem with Nichols and Ulatowski's view is that they do not sufficiently tease out the main components of the two different concepts (Cushman and Mele, in press). They conclude that some concepts of intentional action are desire based and some concepts are belief based. This conclusion is too quick. The chairman cases feature a person who has a belief but lacks a desire to help or harm the environment. This is enough to show that (a) some people think that belief is sufficient bringing about a side effect intentionally, and (b) some people think desire is a necessary condition bringing about a side effect intentionally. However, their scenarios do not test if desire is a sufficient or if belief is a necessary condition for bringing about a side effect

² The asymmetric answer is speculated to be the result of “flexibility” in interpreting “intentionally.”

intentionally. Hence, we cannot conclude that one folk concept is desire based and one folk concept is knowledge based.

Mele and Cushman (2007; Cushman and Mele, in press) also find evidence for multiple concepts while correcting for the shortcoming of Nichols and Ulatowski's study. Using a within-participants design, Mele and Cushman gave participants the chairman scenarios. In addition, participants were given scenarios where the person in the scenario (1) has a belief yet lacks a desire, or (2) has a desire but lacks a belief. Like Nichols and Ulatowski, they found that there are three patterns of responses—some who answer “yes” to both harm and help, some who answer “no” to both, and some who answer “yes” to harm but “no” to help. Because they used additional scenarios featuring (1) and (2), they improve on Nichols and Ulatowski's experiment by filling in the missing conditions. In scenarios where a person has a desire but lacks a belief, Cushman and Mele find that almost all participants think that having a desire is sufficient for acting intentionally. However, participants differ on whether belief is sufficient for acting intentionally. They conclude that there are at least two concepts of intentional action—one that treats belief as a sufficient condition (explain answers of “yes” to harm and help) and one that treats desire as a necessary condition (explains answers of “no” to harm and help).

Cushman and Mele (in press) speculate that there is a third concept which explains the asymmetric answers. Some people may treat desire as a necessary condition for acting intentionally except for morally bad actions. In such cases, they may treat belief as a sufficient condition. Some evidence for this third concept comes from an order effect Cushman and Mele find. They gave participants a total of 16 scenarios, thirteen of which involved morally neutral or morally good side effects. Half of the participants received harm within the first four scenarios

and half received harm in the last five scenarios. When received in the first four scenarios, intentionality ratings were much higher for harm than when it was presented in the last five scenarios. Moreover, participants who answered “no” to harm and help give the same pattern of responses to non-moral cases as those who exhibit the Knobe effect. Cushman and Mele interpret this as providing evidence that when people who otherwise would exhibit the Knobe effect are presented with a series of non-moral or morally good cases involving side effects, they are influenced to think that belief is not a sufficient condition for bringing about a harmful side effect intentionally. Hence, there could be a third concept of intentional action where belief is a sufficient condition only for morally bad actions.

Non-concept Based Explanations

The Trade-Off Hypothesis

Edouard Machery (in press) proposes what he calls the “trade-off hypothesis.” He thinks that most people interpret the chairman cases as ones where the bad side effect is traded for some benefit, and we normally think such trades are done intentionally. Because there is a trade-off—a trade of a desired end along with a foreseen, bad, unintended consequence—and because “we think of costs as being intentionally incurred,” people judge the bad side effect to be brought about intentionally. Cases that involve good side effects are not seen as intentional because there are no associated costs. Thus, the trade-off hypothesis can explain the Knobe effect without referring to our concept of intentional action.

Machery tests the trade-off hypothesis with a pair of cases. In what is supposed to be an analog of the help case, Joe goes into a convenience store to buy a large drink. The cashier tells him that in doing so he gets a commemorative cup. Joe doesn't care at all about the

commemorative cup, he just wants their largest drink. Most people (55%) respond that Joe did not buy the commemorative cup intentionally. In the analog of the harm case, Joe wants to buy a large drink, but the price has gone up by a dollar. He doesn't care at all about the extra dollar, he just wants the largest drink. Ninety-five percent of participants said that he paid the extra dollar intentionally. Hence, in cases where there is a foreseen cost, people judge that it is brought about intentionally. But when there is an unintended, foreseen benefit, people tend to say that the side effect is not brought about intentionally.

The cases that Machery uses are problematic because it is not clear that they involve side effects.³ First, in the help analog, the commemorative cup Joe gets is not a side effect of buying a drink. Getting the commemorative cup is not a *separate* event. Second, in the harm analog, spending an extra dollar is not a side effect of getting a large drink. Joe spends the extra dollar as a means to the desired result, and thereby it is not a *side* effect. Hence, if the trade-off hypothesis is the correct explanation of the Knobe effect, data from different scenarios are required.

Pragmatic Explanations

Adams and Steadman (2004a, 2004b) think the Knobe effect is best explained by conversational implicature. They argue that in harm participants want to say that the chairman is blameworthy. The only way they can express this blame is by saying that the chairman brought about the side effect intentionally. If the participants say that the chairman did not bring about the bad side effect intentionally, then that would conversationally imply that the chairman is not blameworthy for it. So, they use “you did that intentionally” language to assign blame to the chairman for bringing about the bad side effect (Adams and Steadman, 2004a, p. 178). However,

³ To be fair, Machery admits as much.

participants do not want to say the chairman is praiseworthy in help because he did not care at all about bringing about the good side effect; hence, they judge that the chairman did not help the environment intentionally. Because the results are due to conversational implicature, the judgments made about the chairman cases do not tell us anything about the core folk concept of intentional action.

There are some problems with Adams and Steadman's explanation. First, in Knobe's studies, the participants are also asked to assign a praise/blame rating to the chairman. If participants are allowed to express the amount of blame they think the chairman deserves, then the conversational implications should be reduced. After all, participants are allowed to express their ratings of blame so there is no need to conversationally imply that the chairman intentionally brought about the harmful side effect in order to express blame (Malle, 2006). It is unclear why the asymmetry persists in the face of their ability to express blame.

Second, an inference can be made that in other, non-side effects cases, people also use intentionality judgments as a way to conversationally imply blame. If participants use intentionality judgments as a way to blame, then that generates the following prediction: if participants were primed with cases where it is clear that one is blameworthy for something one did unintentionally, then there should be a reduction in intentionality judgments for the blameworthy behavior (Mele, 2003, p. 327). Knobe (2003b) tested this hypothesis. He gave people a case suggested by Mele (2003, p. 327) where a drunk driver is blameworthy for unintentionally killing a family of five. Immediately after the drunk driver case, participants were given either a morally neutral scenario where a person luckily hits a bull's eye in a rifle competition or a morally bad action where a person luckily shoots his aunt in order to inherit a

lot of money. Skill is standardly taken to be a necessary condition for doing something intentionally. So, in both “no skill” cases theory predicts the action is done unintentionally. Most people (96%) thought the drunk driver unintentionally killed the family and the mean response to blame was 5.3 on a 6 point scale (6 being maximum blame). However, after being given the drunk driver case, 84% thought shooting the aunt was done intentionally whereas 40% thought shooting the target was done intentionally (Knobe, 2003b, 317-318). Giving the participants the “debiasing” scenario had no effect on the intentionality judgments. Hence, this is additional evidence that participants are not using intentionality language to assign blame in bad side effect cases. However, it is possible that simply presenting a “debiasing” scenario is not sufficient—perhaps a stronger manipulation is required for participants to really become debiased. For example, participants may have to assume different roles such as a jury member (where one assigns blame) or a psychologist (where one evaluates mental states) [Malle, 2006, p. 104].

Biasing Explanation

Nadelhoffer (2004b) argues the Knobe effect is due to affective biasing. Specifically, Nadelhoffer thinks that the perceived blameworthiness of the chairman fuels the asymmetry. Blameworthiness explains the Knobe effect because harm and help are not analogous. In harm and help, we naturally form negative impressions of the chairman because he does not care about something he should care about. For this reason, people do not want to praise the chairman for bringing about a good side effect, but they do want to blame him for bringing about a bad side effect. Indeed, praise ratings for the chairman intentionally bringing about the good side effect are much lower than blame ratings for the chairman bringing about the bad side effect. Hence, people think that the chairman brings about the harm intentionally but not the help.

This generates the hypothesis that if the cases were truly analogous—if the chairman in the help condition was perceived as praiseworthy—then people would judge the person brings about the side effect intentionally. Nadelhoffer (2004b) tested this hypothesis. His case describes two friends who are competing against each other in an essay contest. Jason helps edit his friend's essay, unconcerned that doing so will lower his own chances of winning the contest; and Jason in fact does not win. Most people (55%) judge that Jason does decrease his chances intentionally and is praiseworthy for doing so. Therefore, when the cases are analogous, the Knobe effect vanishes (Nadelhoffer, 2004b, p. 210).

Nadelhoffer garners additional evidence that blameworthiness influences intentionality ratings. He uses two additional cases that have the same evidential features. In one case, a thief attempts to flee from a policeman. The policeman, in a heroic act to stop the thief, jumps onto the car and is eventually shaken off and killed. The thief doesn't care at all about the policeman, he just wants to get away. In the other version, a carjacker jumps onto the hood of the car to be eventually shaken off and killed. The thief does not care at all about the carjacker, he just wants to get away. Thirty-seven percent of participants thought that the thief intentionally killed the policeman, whereas only 10% thought the thief intentionally killed the carjacker. The mean blame rating in the policeman case was 5.11 on a 6 point scale (1= no blame, 6= a lot of blame) but only 2.01 for the carjacker (Nadelhoffer, 2006b, p. 209). Therefore, the perception of the target of the side effect affects intentionality ratings even when all other evidential features of the case stay the same. Because all the evidential features are the same, Nadelhoffer thinks these results suggest that when “morally loaded features are built into scenarios, these features often trump or override the standard application of the concept of intentional action—thereby

distorting our judgments about intentionality” (2006b, p.213-4).

Malle and Nelson (2003) think that negative affect generated in the chairman cases can explain the Knobe effect. They argue that negative affect can bias one's interpretations of the mental states of another person. They studied the judgments of couples who have fought. Because of the negative affect that is generated in these fights, the parties to the fight think that everything the other person does is intentional—even if it is really not (p. 575). Likewise, because we think that the chairman has done something wrong by harming the environment, which triggers a negative reaction, our judgments about intentionality are biased. Because people see the chairman in a negative light, they are more likely to think that he brings about the harmful side effect intentionally.

The Attention Explanation

Malle (2006) argues that the way in which the scenarios are presented focuses the participants' attention on the evaluative components of the scenarios. The participants are presented with a chairman who does not care at all about the bad side effect. The participants may think that they are supposed to “do” something with this evaluative material, especially because the side effect is so extreme. Because they think they are supposed to do something with this evaluative material, they use it to make “non-technical” intentionality judgments. That is, because they are forced to make an intentionality judgment about the chairman, they use this information to judge that the chairman brought about the side effect intentionally. However, in other contexts where they do not think they are supposed to use this material, the participants would be more inclined to use their core concept and issue a judgment that the chairman did not bring about the bad side effect intentionally.

The Reasons Against Explanation

Keying on a hypothesis by Harman (1976) that one intentionally performs an unintended action if one has a reason not to do it, and one by Mele and Sverdlik (1996) that if one regretfully performs a foreseen side effect, one is likely not to be judged to do it intentionally, Turner (2004) argues that the Knobe effect can be explained by viewing harm as a case where one has a “reason not to perform” the side effect, yet does it anyway. Turner thinks the following conditions are jointly sufficient for a side effect to be brought about intentionally:

- (1) *S* knows that *E* will (or is likely to) occur as a result of *A*-ing
- (2) bringing about *E* counts against *A*-ing (from the *S*'s perspective), and
- (3) *S* does not try to keep *E* from occurring. (Turner, 2004, p. 216-217)

The chairman in harm certainly does not prevent the side effect from occurring and he foresees it coming about. Turner argues that the chairman also thinks that bringing about the side effect is a reason not to perform the action—after all, “implementing the environmentally harmful policy incurs the risk of protest, legal investigation, dissent from the lower echelons of the company, etc.” (2004, p. 217). So, all the conditions are satisfied in harm. However, (2) is not satisfied in help—helping the environment is a reason *for* implementing the program. Hence, the reasons against view can explain the asymmetry because all three conditions are met in harm but not in help.

Knobe (2004a) tested Turner's conditions and found that they are not sufficient to issue intentionality judgments. Knobe's case involves a terrorist who plants a bomb to blow up a nightclub. However, the terrorist discovers that his son is in the nightclub when he plans to blow it up. The only way to save his son is to defuse the bomb, which has the side effect of saving

Americans, whom he hates. The terrorist foresees saving the Americans. For the terrorist, saving the Americans is a reason against defusing the bomb. Because the terrorist defuses the bomb, we can assume that he does not take measures not to save the Americans. So, conditions 1-3 are satisfied. However, Knobe reports that most people judge that saving the Americans is not brought about intentionally (Knobe, 2004a, p. 272). Therefore the conditions offered by Turner are not sufficient for a side effect to be brought about intentionally.

Concluding Remarks

At this point there is no consensus which umbrella view is correct. Indeed, it is possible that both umbrella views are at least in part correct. For example, it may turn out that some groups of people use their concept of intentional action in judging side effects, and other groups of people may not always use their concept of intentional action when judging side effects. Recent research suggests that this possibility is in fact actual (Feltz and Cokely, 2007). In a known order effect with help and harm, Feltz and Cokely found that women, and not men, were affected by the order of presentation. This finding suggests that men and women may use different judgment processes, concepts, or interpretations of the scenarios when they make intentionality judgments.

How are we to determine if there is one concept, multiple concepts, or different judgment processes that generate intentional action intuitions? The current dominant survey methodology in experimental philosophy, while valuable, is ill suited to settle this question. To settle this question, data about the processes that generate these judgments are required. There are several promising avenues of research that provide these data.⁴ One avenue uses stable individual

⁴ Edward Cokely (at the Max Planck Institute for Adaptive Cognitive and Behavior) and I have developed and are currently exploring these different avenues.

differences to help identify judgment processes. As we have seen, there is evidence that groups of people make the same intentionality judgments about the chairman cases. We also know that some individual differences are correlated with different reasoning processes, data encoding, and information retrieval (Cokely, Kelley, and Gilchrist, 2006; Ericsson, Prietula, and Cokely, 2007). If we can identify groups that have stable intentional action intuitions by individual differences, we can begin to better understand what proximal processes are involved in intentionality judgments.

Another avenue of research that provides process level data about intentional action intuitions is talk aloud methods and protocol analysis (Ericsson and Simon, 1980, 1993). For example, we can give participants the chairman cases and ask them to talk aloud as they think about whether the chairman intentionally brings about the side effect. By analyzing these verbal reports, we can understand the processes that generate intentionality judgments. Talk aloud and protocol analysis afford greater fidelity than surveys in determining whether people are subject to biases, have one core concept, have multiple concepts, have different interpretations of scenarios, or have different judgment processes.

Finally, there is a great deal of evidence that many people use fast and frugal heuristics in many domains (Gigerenzer, 1996; Gigerenzer and Goldstein 1996; Gigerenzer and Todd, 1999). So we have good reason to think that there are also heuristics involved in judgments about intentional actions. These intentional action heuristics can be explored, for example, with different manipulations and by measuring reaction times. To illustrate, one heuristic that is presently being explored is that *violations of social norms are intentional*. That is, when someone's behavior deviates from a perceived and socially acceptable norm, we tend to assume

that this behavior is intentional. Identifying intentional action heuristics such as violations of social norms are intentional can help determine what processes lead to intentionality judgments and to the extent that these processes differ. In the end, traveling on all three avenues will go a long way to settle the question of which, or to what extent, each umbrella view is correct.

References

- Adams, F., and Steadman, A. (2004a). Intentional action in ordinary language: Core concept or pragmatic understanding. *Analysis*, 64, 173-181.
- Adams, F., and Steadman, A. (2004b). Intentional actions and moral considerations: Still pragmatic. *Analysis*, 64, 264-267.
- Cokely, E. T., Kelley, C. M., and Gilchrist, A. L. (2006). Sources of individual differences in working memory. *Psychonomic Bulletin and Review*, 13, 991-997.
- Cushman, F., and Mele, A. (in press). Intentional action: Two-and-a-half folk concepts? In J. Knobe and S. Nichols (Eds.), *Experimental Philosophy*. Oxford: Oxford University Press.
- Ericsson, K. A., Prietula, M. J., and Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, 85, 114-121.
- Ericsson, A., and Simon, H. (1993). *Protocol Analysis: Verbal Reports as Data* (Revised ed.). Cambridge: MIT Press.
- Ericsson, K.A., and Simon, H.A. (1980) Verbal reports as data. *Psychological Review*, 87, 215-251.
- Feltz, A., and Cokely, E. (2007). An Anomaly in Intentional Action Ascriptions: More Evidence of Folk Diversity. In D. S. McNamara and J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 1748). Austin, TX: Cognitive Science Society.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: Reply. *Psychological Review*, 103, 592-596.
- Gigerenzer, G., and Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of

- bounded rationality. *Psychological Review*, 103, 650-669.
- Gigerenzer, G., and Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Harman, G. (1976). Practical reasoning. *Review of Metaphysics*, 79, 431-463.
- Knobe, J. (2003a). Intentional action and side-effects in ordinary language. *Analysis*, 63, 190-194.
- Knobe, J. (2003b). Intentional action in folk psychology: an experimental investigation. *Philosophical Psychology*, 16, 309-323.
- Knobe, J. (2004a). Folk psychology and folk morality: Response to critics. *Journal of Theoretical and Philosophical Psychology*, 24, 270-279.
- Knobe, J. (2004b). Intention, intentional action, and moral considerations. *Analysis*, 64, 181-187.
- Knobe, J. (2006). The concept of intentional action: A case study in uses of folk psychology. *Philosophical Studies*, 130, 203-231.
- Knobe, J., and Burra, A. (2006). The folk concept of intention and intentional action: A cross-cultural study. *Journal of Culture and Cognition*, 6, 113-132.
- Knobe, J., and Mendlow, G. (2004). The good, the bad, and the blameworthy: Understanding the role of evaluative reasoning in folk psychology. *Journal of Theoretical and Philosophical Psychology*, 24, 252-258.
- Leslie, A., Knobe, J., and Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science*, 17, 421-427.
- Machery, E. (in press). Understanding the folk concept of intentional action: Philosophical and experimental issues. *Mind and Language*.

- Malle, B. (2006). Intentionality, morality, and their relationship in human judgment. *Journal of Cognition and Culture*, 6, 87-112.
- Malle, B., and Nelson, S. (2003). Judging *mens rea*: The tension between folk concepts and legal concepts of intentionality. *Behavioral Sciences and the Law*, 21, 563-580.
- McCann, H. (2005). Intentional action and intending: Recent empirical studies. *Philosophical Psychology*, 18(6), 737/748.
- Mele, A. (2003). Intentional action: Controversies, data, and core hypotheses. *Philosophical Psychology*, 16, 325-340.
- Mele, A. (2001). Acting intentionally: Probing folk intuitions. In B. F. Malle, L.J. Moses, and D.A. Baldwin (Eds.), *Intentions and Intentionality* (pp. 27-43). Cambridge, MA, MIT Press.
- Mele, A. (1992). Recent work on intentional action. *American Philosophical Quarterly*, 29, 199-217.
- Mele, A., and Cushman, F. (2007). Intentional action, folk judgments, and stories: Sorting things out. *Midwest Studies in Philosophy*, 31, 184-201.
- Mele, A., and Sverdlik, S. (1996). Intention, intentional action, and moral responsibility. *Philosophical Studies*, 82, 265-287.
- Nadelhoffer, T. (2004a). The Butler problem revisited. *Analysis*, 64, 277-284.
- Nadelhoffer, T. (2004b). On praise, side effects, and folk ascriptions of intentional action. *Journal of Theoretical and Philosophical Psychology*, 24, 196-213.
- Nadelhoffer, T. (2004c). Blame, badness, and intentional action: A reply to Knobe and Mendlow. *Journal of Theoretical and Philosophical Psychology*, 24, 259-269.

- Nadelhoffer, T. (2005). Skill, luck, control, and intentional action. *Philosophical Psychology*, 18, 341-352.
- Nadelhoffer, T. (2006a). Desire, foresight, intentions, and intentional action: Probing folk intuitions. *Journal of Cognition and Culture*, 6, 133-157.
- Nadelhoffer, T. (2006b). Bad acts, blameworthy agents, and intentional actions: Some problems for juror impartiality. *Philosophical Explorations*, 9, 203-220.
- Nadelhoffer, T. (2006c). On trying to save the simple view. *Mind and Language*, 21, 565-586.
- Nadelhoffer, T., and Nahmias, E. (2007). The past and future of experimental philosophy. *Philosophical Explorations*, 10, 123-149.
- Nichols, S., and Ulatowski, J. (2007). Intuitions and Individual Differences: The Knobe Effect Revisited. *Mind and Language*, 22, 346-365.
- Phelan, M., and Sarkissian, H. (in press). The folk strike back: Or, why you didn't do it intentionally, though it was bad and you knew it. *Philosophical Studies*.
- Sverdlik, S. (2004). Intentionality and moral judgments in commonsense thought about action. *Journal of Theoretical and Philosophical Psychology*, 34, 224-236.
- Turner, J. (2004). Folk intuitions, asymmetry, and intentional side effects. *Journal of Theoretical and Philosophical Psychology*, 24, 214-219.