

Intentional action:

Controversies, data, and core hypotheses

Alfred R. Mele

Florida State University

Department of Philosophy

Tallahassee, FL 32306-1500 USA

almele@mailier.fsu.edu

ABSTRACT This article reviews some recent empirical work on lay judgments about what agents do intentionally and intend in various stories and explores its bearing on the philosophical project of providing a conceptual analysis of intentional action. The article is a case study of the potential bearing of empirical studies of a variety of folk concepts on philosophical efforts to analyze those concepts and vice versa. Topics examined include double effect; the influence of moral considerations on judgments about what is done intentionally and about what is intended; the influence of considerations of luck, skill, and causal deviance on judgments about what agents do intentionally; what interesting properties all cases of intentional action might share; and the debate between proponents of, respectively, “the Simple View” of the connection between intentional action and intention and “the Single Phenomenon View” of that connection.

A substantial body of literature is devoted to the project of analyzing intentional action.¹ In this article, I explore the bearing on that project of some recent empirical work on lay judgments about what is done intentionally and about what is intended. This article may reasonably be regarded as a case study of the potential bearing of empirical studies of a range of folk concepts on philosophical efforts to analyze those concepts and, likewise, of the potential bearing of attempted philosophical analyses of folk concepts on empirical studies of those concepts.

1. Two disputes and some empirical studies

Philosophers have disagreed about the relevance of considerations of moral responsibility to correct judgments about what is and is not done intentionally. Michael Bratman has argued that to do justice both to “our concern with responsibility” and to regularities of a kind required if intentions are to play explanatory roles plausibly attributed to them, “we need to allow our concern with responsibility to shape what is done intentionally without similarly shaping what is intended” (1987, p. 125). Steven Sverdlik and I have defended the opposing view that although Bratman is right in holding that correct attributions of intentions have no special sensitivity to moral responsibility, correct judgments about what is done intentionally also are not sensitive to moral responsibility (Mele & Sverdlik 1996).

A related disagreement concerns actions that are side-effects of intended intentional actions. Gilbert Harman asks his readers to suppose that “in firing his gun,” a sniper who is trying to kill a soldier “knowingly alerts the enemy to his presence” (1976, p. 433). He claims that although the sniper “does not intend to alert the enemy,” he *intentionally* alerts the enemy, “thinking that the gain is worth the possible cost.” Bratman makes a similar claim about a runner who reluctantly wears down some heirloom shoes (1987, p. 123; cf. Ginet 1990, pp. 75-6). In

Mele & Sverdlik 1996, an opposing view of such “side-effect actions” is defended. Because Harman’s sniper and Bratman’s runner do not unknowingly or accidentally perform the actions at issue, perhaps most people would deny that the sniper *unintentionally* alerted the enemy and that the runner *unintentionally* wore down his shoes. But that denial would not obviously commit them to holding that these actions are *intentional*. There may be a middle ground between intentional and unintentional action. Possibly, actions that an agent in no way aims at performing and are not parts of anything he is aiming at but also are not performed unknowingly or accidentally are properly located on that middle ground. They may be *nonintentional*, as opposed to *unintentional*. The view that this is so is motivated in Mele & Sverdlik 1996.

There is empirical evidence that Bratman and Harman are more sensitive to the folk concept of intentional action than are Sverdlik and I, for example. In some recent studies, Joshua Knobe produces evidence for the propositions that commonsense judgments about what is done intentionally are strongly influenced by moral considerations and that some side-effect actions are commonly deemed intentional. I will summarize these studies shortly. First, a brief comment is in order on Knobe’s methodology and my use of the expression “the folk concept of intentional action.” It is an interesting question whether every culture has a (or the) concept of intentional action, but Knobe’s studies do not answer it. They are conducted with people spending leisure time in Manhattan parks – multicultural places to be sure, but not multicultural enough for this purpose. By “the folk concept of intentional action” here I mean a concept of intentional action that I will suppose to be reflected in the judgments a substantial majority of Knobe’s subjects make about stories they are given – judgments that certain actions described in those stories are intentional, unintentional, or not intentional.²

In a study of lay intentionality judgments about morally wrong and morally neutral actions involving considerable luck, Knobe finds that people respond very differently to the following two cases (2003, p. 000):

1. Jake desperately wants to win the rifle contest. He knows that he will only win the contest if he hits the bull's-eye. He raises the rifle, gets the bull's-eye in his sights, and presses the trigger. But Jake isn't very good at using his rifle. His hand slips on the barrel of the gun, and the shot goes wild. Nonetheless, the bullet lands directly on the bull's-eye. Jake wins the contest.
2. Jake desperately wants to have more money. He knows that he will inherit a lot of money when his aunt dies. One day he sees his aunt walking by the window. He raises his rifle, gets her in his sights, and presses the trigger. But Jake isn't very good at using his rifle. His hand slips on the barrel of the gun, and the shot goes wild. Nonetheless, the bullet hits her directly in the heart. She dies instantly.

Knobe reports that 28% of the 18 respondents to case 1 said that Jake intentionally hit the bull's-eye whereas 76% of the 21 respondents to case 2 said that Jake intentionally killed his aunt.

Obviously, the main difference between the cases is a moral one.³

Knobe's result is confirmation for my conjecture elsewhere that a case like 2 would get "a significantly higher intentionality rating" from untutored respondents than a case like 1 (Mele 2001, p. 40). I also conjectured, however, that with a group of subjects to whom it had recently been made salient that people can be blameworthy for things they do unintentionally (e.g., that a drunk driver who loses control of his car and unintentionally kills a family of five can be

blameworthy for killing them), the difference in intentionality ratings of cases like 1 and 2 would shrink significantly (p. 41). Knobe tests the latter conjecture too. He reports that the result of an experiment with people primed in this way is very similar to the original result: 84% of the 32 respondents said that Jake intentionally killed his aunt in case 2 and 40% of the 25 respondents said that he intentionally hit the bull's-eye in case 1 (2003, p. 000).

In another study, Knobe investigates common-sense judgments about side-effect actions with moral significance (n.d.-a). Subjects respond very differently to the following two cases:

3. 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 . . . 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.

4. 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, and it will also help the environment." The chairman . . . 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 reports that 82% of the 38 respondents to case 3 said that the chairman intentionally harmed the environment whereas 77% of the of the 40 respondents to case 4 said that the chairman did *not* intentionally help the environment. Evidently, respondents' reactions vary with their moral assessment of the side-effect action. Knobe also asked respondents to case 3 to

indicate how blameworthy the chairman is on a scale from 0 to 6 and asked respondents to case 4 to indicate how praiseworthy he is on the same scale. The mean rating for blame was 5.2 whereas that for praise was 2.1. Knobe suggests that this asymmetry may be a source of the asymmetry in their intentionality judgments.

2. A shortage of data

Tests of the sort Knobe has conducted provide evidence about how the groups he surveys understand what it is to do something intentionally. We need more data of that kind to be in a position reliably to formulate the folk concept of intentional action in terms of necessary and sufficient conditions. This need is the primary topic of this section. I discuss it for two reasons. First, it helps to explain my procedure in this article: it explains, for example, why I do not try to produce an analysis of the folk concept of intentional action. Second, I hope to motivate tests of folk judgments about cases of the kind I sketch.

Harman offers a diagnosis of the alleged intuition that “the sniper intentionally alerts the enemy to his presence”: “We say this because the sniper acts in the face of a reason not to alert the enemy to his presence” (1976, p. 434). He adds:

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. (p. 434)

Part of Harman’s purpose in making the latter claim is to account for the difference between a pair of alleged intuitions: (1) “at the firing range a sniper intentionally shoots a bull’s-eye only if

that is something he can do at will”; and (2) a sniper who is trying to kill a soldier kills him intentionally even if his chance of killing him is no greater than the chance he would have of hitting a certain bull’s-eye and the latter chance is such that, if he were to hit the bull’s eye, he would not be said to hit it intentionally (pp. 443-44). Harman contends that

the reason why we say that the sniper intentionally kills the soldier but do not say that he intentionally shoots a bull’s-eye is that we think that there is something wrong with killing and nothing wrong with shooting a bull’s-eye. (p. 434)

The pair of intuitions just reported, 1 and 2, are in line with the responses Knobe’s subjects make to the two Jake cases, as is the sort of diagnosis Harman offers of the difference between them. (Whether most people would require the ability to hit a bull’s-eye “at will” or something more modest for an intentional hitting of it is a separable question.) However, in light of Knobe’s data, Harman’s remarks prompt several questions. Would only the difference between agents’ having and lacking *moral* reasons for not doing something they try to do and succeed in doing generate asymmetrical judgments of the kind made about the Jake cases or would the difference between having and lacking *nonmoral* reasons for refraining have a similar effect? Suppose, for example, that the following changes are made to case 1. Jake, a man of modest means, knows that he can win the contest either by hitting the bull’s-eye or by coming closer to hitting it than his opponent does, that the prize for the winner is \$750, and that because this bull’s-eye is an expensive device, whoever hits it must pay \$500 to replace it. (As in case 1, Jake aims at the bull’s-eye: he believes his opponent is very skillful and might well hit the bull’s-eye or come extremely close.) Would this have an effect on folk judgments about whether Jake

hit the bull's-eye intentionally? Also, would most respondents say that Jake intentionally destroyed the bull's-eye?

Turn to the sniper. Suppose it is made clear that his alerting the enemy would endanger only himself and that his plan is not at all reckless, so that respondents will be disinclined to see the fact that he would alert the enemy by firing as a *moral* reason not to fire. Would the sniper's alerting the enemy get as high an intentionality rating as the chairman's harming the environment? Imagine that lay folk are presented with a scenario in which a sniper knows that he would alert only his comrades to his presence by firing at a solitary enemy soldier and that this would be good for him but does not fire even partly for that reason. He fires only to eliminate the enemy soldier and with no concern to promote his own welfare. Would this sniper's alerting his comrades get an intentionality rating close to that of the chairman's helping the environment? Would it get pretty much the same rating as the other sniper's alerting the enemy?

Until tests are run, any answers offered to the questions raised in the preceding two paragraphs would be largely speculative. Tests would help us learn whether nonmoral reasons against doing what one does – and nonmoral reasons for doing what one does that do not influence the agent – sometimes have roughly the same impact as moral reasons on lay judgments about what is and is not done intentionally.⁴

Other kinds of scenario also require investigation. Here is a case featuring a side-effect that the agent unsuccessfully tries to avoid. In this, it differs from the sniper and chairman cases.

5. Pat is dying. His doctor, Doris, judges both that death is very likely and that an operation has some chance of saving his life. She also knows that the operation itself has a good chance of killing Pat. Doris decides to perform the operation, and she takes great

care to minimize the chance that it will kill Pat. Unfortunately, her efforts are unsuccessful. The operation proves fatal. (cf. Mele & Sverdlik 1996, p. 278)

Doris has a reason not to perform the operation. Would majority opinion be that Doris intentionally killed Pat? My guess is that the answer is *no*, but I have been wrong about such things in the past.

Cases of lucky success that differ in a way to be explained from Jake's lucky killing of his aunt need attention too. Consider the following story:

6. In a situation reminiscent of James Bond movies, an assassin who is standing at a gaming table in an abandoned casino wishes to kill an innocent countess within the next few minutes. He realizes that the only way he can do this is by throwing a six with the ordinary six-sided die in his pocket. The countess is locked in a room in another building. There is a bomb in her room, and the assassin can detonate it only by producing a six-dotted image on the lens of a camera that is focused on the top of the table and wired to the bomb. The camera is so positioned that it will register only the *top* face of the die. Approaching the table (to place the die six-up on it, for example, or to retrieve his die after throwing it) would deactivate the apparatus. Although he has no special powers over dice, the assassin does, of course, have a one-in-six chance of tossing a six. Knowing all this, he tosses the die onto the table, hoping that it will land six-up. The die lands six-up. By throwing a six, the assassin detonates the bomb, thereby killing the countess. (cf. Mele & Sverdlik 1996, p. 279)

This story is similar to Knobe's story about Jake's killing his aunt in that considerable luck is involved. However, there is an interesting difference. It is common knowledge that people differ greatly in their skill at hitting what they are shooting at, and it is uncontroversial that many people can intentionally hit certain targets. But it also is common knowledge that whether people roll a six when they roll a fair die is not at all a matter of skill at rolling sixes: it is just a matter of chance. Would a substantial majority of respondents say that the assassin intentionally killed the countess? If so, would a substantial majority who are asked only whether the assassin intentionally rolled a six say that he did? If the answer to the latter question is yes, that is evidence that the folk-concept of intentional action does not oppose intentional action to mere chance nearly as strongly as many philosophers of action do (see Butler 1978). Also, should it turn out that the majority of respondents would *not* say that the assassin intentionally rolled a six even though a substantial majority would say that he intentionally killed the countess, that would be evidence that the folk concept of intentional action allows that agents sometimes intentionally do *A* without intentionally doing *B* even when they do *A* precisely by doing *B* and doing *B* was a highly salient, necessary part of their strategy for doing *A*. (The assertion that agents sometimes intentionally do *A* without intentionally doing *B* under such circumstances would, I conjecture, strike most philosophers of action, at least initially, as counterintuitive.)

Here is a standard case of causal deviance.

7. Carlo tries to kill his aunt by running her down with his car. His attempt so unnerves him that he loses control of the car. It careens off of several other vehicles, some parking meters, and a lamp post before it crashes into his aunt. She dies within seconds of impact. (cf. Chisholm 1966, pp. 19-20)

Would a majority of lay respondents say that Carlo's killing his aunt is an intentional action?

Knobe reports that 96% of lay respondents to a story in which a drunk driver loses control of his car and kills a family of five say that the driver unintentionally kills the family (2003, p. 000).

But there is a difference between that case and case 7: the drunk driver is not trying to kill anyone when he loses control, and Carlo is. Possibly, this difference would make a great difference in lay intentionality judgments, and possibly not. One might suppose that lay reactions to Carlo's case would be very similar to the reported reactions to case 2, in which Jake kills his aunt with a lucky shot. However, it is not obvious that this would be so. Differences between the cases might influence folk judgments. The difference in salient intervening events between the pertinent controlled movements and the bad consequences in the two cases may have an effect. Also, whereas it is clear that Carlo completely loses control of his car, it is not clear that Jake completely loses control of his rifle. That related difference may influence folk judgments.

Clearly, a great deal remains to be done in empirical tests of the folk concept of intentional action. Until it is done, theorists will be in no position confidently to construct a full analysis of that concept.

3. Knobe's data and "the Simple View"

How should philosophers of action respond to the results Knobe has produced? In the present section I focus this question by setting it in the context of a specific philosophical controversy. In Section 1, I mentioned a pair of philosophical disagreements to which Knobe's results are relevant, and I will return to them later. The controversy featured in the present section is over

what Bratman has dubbed “the Simple View” (1984, p. 377), the thesis that intentionally *A*-ing entails intending to *A*.

In Section 1, I reported that 82% of Knobe’s (n.d.-a) respondents to case 3 said that the chairman intentionally harmed the environment. In a subsequent study (n.d.-b), Knobe asked some people whether the chairman intentionally harmed the environment and asked others whether he intended to harm it. 87% of the 15 respondents to the former question said that he intentionally did this whereas only 29% of the 17 respondents to the latter question said that he intended to do it. This result is at odds with the Simple View. (Incidentally, Knobe asked parallel questions about the chairman in case 4 who benefited the environment without caring whether he did. 20% said that he intentionally benefited the environment and none said that he intended to benefit it.)

Frederick Adams (1986) and Hugh McCann (1986, 1991) have vigorously defended the Simple View against objections. Adams asserts that “we rely (at least implicitly) on the Simple View to distinguish” intentional from unintentional actions (1986, p. 285). McCann claims that “we get at least the beginning of a correct distinction” on this topic if we endorse the Simple View (1986, pp. 191-92) and that distinguishing between nonintended intentional actions and nonintended nonintentional ones “promises to be a most intractable problem” (p. 192). However, Knobe’s subjects seem not to be relying on the Simple View when they judge that the chairman intentionally harms the environment.

How should advocates of the Simple View respond to Knobe’s data? They might challenge his methods or refuse to consider his results until follow-up studies are done. A positive suggestion they might make is that Socratic questioning would bring people’s judgments into line with the Simple View. Someone conducting such questioning might ask a group of lay

folk whether the chairman in case 3 intentionally harmed the environment and whether he intended to harm it and ask parallel questions about case 4. If many people who say that the chairman intentionally harmed the environment deny that he intended to harm it and say as well that the chairman in case 4 neither intended to help the environment nor helped it intentionally, the questioner might try to lead those folks to the view that their judgments are inconsistent and then see how they resolve the inconsistency. For example, the Socratic figure might try to persuade respondents that it is inconsistent to believe both that the chairman who harmed the environment without caring about that harmed it intentionally and that the chairman who helped the environment without caring about that did not help it intentionally. Obviously, if there is a logical inconsistency involving the contents of these beliefs, it includes an additional item. The following would turn the trick: (*S*) Because when any two actions occur in exactly similar ways and their agents have exactly similar abilities and exactly similar attitudes toward their respective relevant prospective actions either both actions are intentional or neither is, any two such actions that differ in that one is a harming and the other is a helping are either both intentional or both not. (Of course, the fan of the Simple View would hope that if respondents continue to say that the chairmen did not intend to harm/help the environment, they would move to the view that the chairman who harms the environment does not intentionally harm it.)

If vigorous proponents of the Simple View – Adams and McCann, for example – were to try to persuade a majority of patient people spending leisure time in Manhattan parks that *S* is true, would they succeed? I do not know. But it is arguable that even if they were to succeed, the responses Knobe gathered would reveal more about the folk concept of intentional action than responses made by people Adams and McCann persuade to affirm *S*. *S* might appear plausible to the imagined lay respondents partly because of how they understand intentional action but also

partly because they are outmatched by their questioners. Lay folk may not see how to defend their asymmetrical judgments in the face of a challenge to produce a unified account of intentional action, and they may be impressed that articulate proponents of the Simple View seem to be on their way to a relatively simple and straightforward theory about the difference between intentional and nonintentional actions.

I mentioned McCann's claim that distinguishing between nonintended intentional actions and nonintended nonintentional ones is an intractable problem. On the present score at least, a spokesperson for lay folk can say, "Look, my concept of intentional action has a normative dimension of such a kind that the chairman who harms the environment in case 3 intentionally harms it whereas the chairman who helps the environment in case 4 does not intentionally help it. I'm not being inconsistent in making these judgments unless I also accept a thesis like *S*, but I reject *S*. I reject it because it flies in the face of my confident judgments about cases." McCann writes approvingly of "preanalytic data" (1991, pp. 31, 34). If there are preanalytic judgments on the topics at issue, lay judgments about what is and is not done intentionally in Knobe's stories and about what is and is not intended in those stories are likely to be significantly more preanalytic than McCann's intuitions – and mine too, of course – about these things. And although making the distinction at issue given McCann's intuitions about cases may be impossible, it may be that the intuitions of a great majority of lay folk can be captured in an account of the difference between nonintended intentional actions and nonintended nonintentional ones. Of course, in light of what I said in Section 2, I believe that philosophers are not yet in a position reliably to undertake the capturing task.

McCann says that "the Simple View . . . pertains to the everyday concept of intending, not a stipulated one" (1991, p. 33). Presumably, he would have said the same, if asked, about the

concept of intentional action to which the view pertains. Some of Knobe's data give McCann good reason to doubt that his concept of intentional action is a folk concept of such action.

I mentioned that a Socratic questioner might, unlike Knobe, ask the *same* people about cases 3 and 4 and ask them both whether the chairman intentionally harmed/helped the environment and whether he intended to do so. Recall that Knobe's subjects are each asked one of the two questions about one of the two cases, with the following results:

- a.* (Case 3) The chairman intentionally harmed the environment: 82% (87% in another study).
- b.* (Case 4) The chairman did *not* intentionally help the environment: 77%.
- c.* (Case 3) The chairman intended to harm the environment: 29%.
- d.* (Case 4) The chairman intended to benefit the environment: 0%.

Suppose that when a single group is asked all four questions, the gap between *a* and *b* is much smaller and the same is true of the gap between what is said to be done intentionally and what is said to be intended. What would this indicate?

One answer is that the opportunity to reflect on all four questions enables subjects to think more carefully about relevant issues and to give more accurate responses. But there is an alternative interpretation of the imagined result. One might distinguish between folk concepts and folk theories (collections of theoretical beliefs). It may be suggested that lay folk are disposed to assent to the abstract theoretical proposition that people intentionally *A* only if they intend to *A*, that this disposition would be activated by their being asked both whether the chairman intentionally harmed the environment and whether he intended to harm it (that is, by

their being asked that *pair* of questions), and that the disposition has less to do with the folk concept of intentional action than with folk theories about intentional action (see Mele 2001, pp. 27-35). The suggestion, in short, is that Knobe's methods get at the folk concept better than the imagined battery-of-questions method would. If and when the battery method is used, depending on the results, these issues may require careful attention.

4. Knobe's data and core intentional action

At one time, as I mentioned, I thought that although lay folk would make asymmetrical intentionality judgments about pairs of cases like 1 and 2 (the shaky gunman cases), the asymmetry would be much smaller if they were primed with the recognition that people can be morally responsible for things they do unintentionally. My hypothesis, roughly, was that their judgments about what is and is not done intentionally would sometimes be corrupted by a combination of the warranted belief that the agent in the moral member of the pair of otherwise symmetrical cases is blameworthy for *A*-ing and the unreflective unwarranted assumption that because he is blameworthy for that, he *A*-ed intentionally. Knobe's finding that groups of respondents to whom it has just been made salient that people can be blameworthy for things they do unintentionally display no less asymmetry in their judgments about these cases is bad news for that hypothesis of mine.

If I had been asked back then to describe my main interest in the philosophy of action, I would have said something like this: to construct a viable theory about how agents produce their intentional actions. Such a theory would involve explanatory items, including, in my view, events of intention acquisition, the persistence of intentions, and the roles these things play in initiating, guiding, and sustaining some actions (see Mele 1992a and 2003, ch. 2). It would also

involve a view about the nature of the product at issue: intentional action. After all, someone who lacks a view of what a certain phenomenon is will have a hard time thinking about how the instances of that phenomenon are produced. I could, in good conscience, hold on to exactly this conception of things if I could persuasively argue that Knobe's study with the primed subjects is seriously flawed or that folk judgments about intentional action are seriously confused. But the truth, although I did not realize it then, is that my main interest was to construct a viable theory about how agents produce their intentional actions, *as I* (and many philosophers of action, I believe) *conceived of intentional actions*. Describing my project this way would certainly seem to narrow its potential interest. However, the constriction may be reversible. Suppose it were shown that what I and some other philosophers were thinking of as intentional action lies at the core of a much broader folk concept of intentional action – that, rather than being an invention of philosophers, it is a philosophical extraction of something central to the folk concept. This anchor in the folk concept would give the topic a broader significance than it would otherwise have.

A philosopher might try to construct a theory about how agents produce their intentional actions as lay folk conceive of such actions. Such a theorist may find that the philosophically most interesting work is focused on a core region of the space marked out by the folk concept. Imagine explaining how shaky Jake produced his intentional action of killing his aunt in case 2. One might apply a view of the production of “core” intentional actions to Jake's aiming the rifle at his aunt and his pulling the trigger and then just describe the actual causal chain linking these actions to his aunt's death. No special ingenuity is called for in the latter connection. Of course, there is the question how far the folk concept of intentional action extends in cases of this kind. Suppose, for example, that just after he aims his rifle for the purpose of killing his aunt, Jake

sneezes in a way that causes his trigger finger to jerk. The rifle goes off, and his aunt is killed. If the majority of lay folk count this as an intentional killing, so be it. An account of the production of this intentional action may be an account of the “core” intentional action – Jake’s aiming his rifle at his aunt – plus a description of the causal chain linking that action to the poor woman’s death. The same approach may be taken to the killings in cases 6 and 7 (the die and uncontrolled car cases), if they are counted as intentional. By the same token, if most lay folk count none – or only some – of these killings as intentional, it may be possible to accommodate that fact. The optimal strategy for a philosopher with the project under discussion may be to produce a theory of the production of core intentional actions, wait to learn how far the folk concept extends, and then formulate general conditions under which actions outside the core that are generated (see Goldman 1970) by actions inside the core count as intentional actions.

The success of this final part of the strategy – formulating the general conditions – depends on folk judgments’ being suitably patterned. If they are not, arguments that the folk concept is incoherent and needs to be replaced will emerge. Such arguments would be expected from proponents of the Simple View, for example. Until further empirical tests are done, we will be in no position to make reliable judgments about the coherence of the folk concept.

Elsewhere, I have voiced the platitude that

if there is a widely shared concept of intentional action . . . a philosophical analysis of intentional action that is wholly unconstrained by that concept runs the risk of having nothing more than a philosophical fiction as its subject matter. (2001, p. 27).

How much does a central and relatively pure kind of intentional action that is a prime candidate for being explained, in relatively standard ways, in terms of initiation, sustaining, and guidance by intentions have in common with the range of actions that the folk concept counts as intentional? This is an interesting question, but one moral of Section 2 is that we are not yet in a position confidently to answer it. Further testing is required.

Partly in light of Knobe's data, it may be thought that philosophers in the business of developing theories about how intentional actions are produced would do best to concentrate exclusively on how attempts – tryings – are produced. It is plausible that trying to *A* is essentially intentional (see McCann 1986, p. 201), and there is very high agreement among Knobe's subjects about what agents are trying to do in parallel moral and nonmoral cases.⁵ Now, philosophers of action have been much more ambitious than this. Compare case 7 with a case in which an agent, Conrad, retains control of his car in his attempt to kill his aunt and drives directly into her, as planned. She dies shortly after impact. Some philosophers have tried to produce accounts of the production of intentional actions that would accommodate not only the *attempts* in these two cases but also Conrad's driving into his aunt and his killing his aunt. Even if the majority of lay folk would say that Carlo intentionally killed his aunt in case 7, they would have to grant that there is a clear difference between the ways in which the two killings happen. A flexible philosopher of action may say that he or she is concerned to produce an account of the production of actions that are intentional in a certain important way and then proceed to motivate the project.

If the folk concept of intentional action is as broad as Knobe's research suggests it may be, how might a philosopher go about motivating the idea that it includes a central or core element of special importance? One approach is to look for interesting properties that all *cases* of

intentional action have in common even if not all intentional actions have them. To explain this last remark, I need to say something about the individuation of actions.

Donald Davidson invites his readers to consider a story in which “I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home” (1980, p. 4). How many actions does the agent, Don, perform? Davidson’s *coarse-grained* view of action individuation yields the answer that he performs one action “of which four descriptions have been given” (1980, p. 4; cf. Anscombe 1963). An alternative *fine-grained* view treats *A* and *B* as different actions if, in performing them, the agent exemplifies different act-properties (Goldman 1970). On this view, Don performs at least four actions, since the act-properties at issue are distinct. An agent may exemplify any of these act-properties without exemplifying any of the others. One may even turn on a light in a room without illuminating the room: the light may be painted black. Another alternative, a componential view, represents Don’s illuminating the room as an action having various components, including (but not limited to) his moving his arm, his flipping the switch, and the light’s going on (Ginet 1990, Thalberg 1977, Thomson 1977). Where proponents of the coarse-grained and fine-grained theories find, respectively, a single action under different descriptions and a collection of intimately related actions, advocates of the various componential views locate a larger action having smaller actions among its parts.

In this article, I have been proceeding in a neutral way regarding these competing theories of action individuation, and I will continue to do so. Advocates of the coarse-grained view should understand my action labels (e.g., “Jake’s killing his aunt”) as labels for actions under descriptions rather than for actions themselves, and they should understand me to mean “an action under a description” by “an action.”

With this background in place, the idea that all *cases* of intentional action have something interesting in common even if not all intentional actions have it in common is easily explained. For example, it may be true that in every case of intentional action, the agent does something that she intends to do even if it is false that every intentional action is intended.⁶ The case of the chairman who harms the environment without caring whether he does may be used to illustrate this suggestion. The chairman intends to implement the program even if he does not intend to harm the environment. And even if he intentionally harms the environment without intending to harm it, his harming it is a side-effect of an *intended* intentional action of his.

Here is a hypothesis: (*HI*) Setting aside so-called “negative actions” (e.g., not voting in an election or not teaching one’s class, when one does not make an effort not to do those things), the folk concept of intentional action does not permit cases of intentional action in which the agent does not perform any intended intentional actions.⁷ One can test this hypothesis using Knobe’s methods, and I look forward to seeing the results of such tests. If the hypothesis is true, its truth speaks in favor of the idea that intended intentional action lies at the core of the folk concept of intentional action. Obviously, moral properties of actions will not lie at the core of that concept, if a property’s lying at the core of the concept is a matter of its being a property that no case of intentional action can lack. For many cases of intentional action are nonmoral. Nor will the existence of a significant reason – moral or otherwise – for an agent’s not doing what he does lie at the core of the concept in this sense. For there are cases of intentional action in which there is no such reason. (I intentionally opened my door this morning, and there was no significant reason for my not opening it.)⁸

A more exacting hypothesis also merits consideration. It may be that intended intentional actions of a certain kind or kinds lie at the core of the folk concept of intentional action while

intended intentional actions of other kinds do not. In the scenario in which shaky Jake kills his aunt with a lucky shot (case 2), most of Knobe's respondents deem the killing intentional. Most people might deem it intended as well. Undoubtedly, the great majority of Knobe's respondents would also say that Jake intentionally pulled the trigger, for example. Now, if both the trigger pulling and the killing are intentional, the former is intentional in a way in which the latter is not. When the moral features are stripped away from Jake's case, as in the bull's-eye variant, popular opinion has it that his hitting what he is shooting at is not an intentional action, but the common-sense judgment that his pulling the trigger is an intentional action would not change.⁹ His pulling the trigger is deemed intentional in a more fundamental way, a way that is independent of the moral status of what he is up to in pulling it. I need a name for any intended intentional actions performed in nonmoral cases of action and any intended intentional actions performed in moral cases of action that the majority of lay folk would count as intentional even if the moral aspect of the cases were stripped away. I call them intended *C-intentional* actions. A more exacting hypothesis than *H1* may be formulated as follows: (*H2*) Setting aside negative actions, the folk concept of intentional action does not permit cases of intentional action in which the agent does not perform any intended C-intentional actions.

Here is a hypothesis of another kind: Setting aside negative actions, an adequate theory of the production of intentional actions will have at its core an adequate theory of the production of a core kind of intended intentional actions – either intended C-intentional actions or something narrower. A narrower kind of action will be called for if the difference between there being and there not being *nonmoral* reasons for agents' not doing what they are trying to do systematically influences folk judgments in such a way that a substantial majority of respondents count successful attempts in some "opposing reason" cases as intentional actions whereas a substantial

majority do not count successful attempts in parallel “no opposing reason” cases as intentional actions. (Of course, whether the difference at issue makes a difference in intentionality judgments may depend on the weight of the nonmoral reason.) This hypothesis is in trouble if lay opinion backs the following two propositions about some cases of non-negative action: the agent performs no intended intentional actions (or no intended C-intentional actions), and she does perform an intentional action.¹⁰ Whether cases can be constructed that generate this response remains to be seen.¹¹

5. Conclusion

Philosophers interested in analyzing folk concepts certainly should encourage work of the sort Knobe has done. Lay responses to questions about what is or is not done intentionally in well-conceived stories provide a guide to the folk concept of intentional action. Information about other folk concepts – for example, folk concepts of knowledge, free choice, causation, and moral responsibility – may be gathered in the same general way. Of course, there is the question how much philosophers should care about some folk concepts that have the same name, as it were, as concepts in which they have a special interest. There is no reason to think that the answer will be the same across the board. Some philosophers interested in the nature of space or time, for example, may have little incentive to study folk concepts of these things. However, in some spheres there are philosophers who say that their concern is an “everyday” – or folk – concept of X and not something that their philosophical opponents have cooked up. For example, McCann (1991, p. 33) says this about intention, as I mentioned; T. M. Scanlon says something similar about desire (1998, p. 40); and the question whether the folk concepts of free will and moral responsibility admit of instantiation is a topic of debate (see, e.g., Strawson 1994).

How should we greet the predictable suggestion that the folk concept of intentional action lies in a reflective equilibrium between folk judgments about cases and proposed philosophical analyses of intentional action? My suggestion is that we should wait and see whether comprehensive testing of lay folk yields a consistent collection of majority judgments about cases. If the set is consistent, we can generate an analysis from the data. If the set is inconsistent, a move toward reflective equilibrium may be in order. As John Rawls understands the process, lay folk would individually reach reflective equilibrium about what intentional action is after they have “weighed various proposed conceptions” of intentional action and have “either revised [their] judgments to accord with one of them or held fast to [their] initial convictions (and the corresponding conception)” (1971, p. 48). In effect, this turns lay folk into philosophy students; and unless they hold “fast to [their] initial convictions,” the concept lying in reflective equilibrium may be a hybrid concept rather than a folk concept. This is not to deny that a hybrid concept of intentional action would be useful for some purposes. Of course, the purposes would need to be specified and the alleged utility would need to be defended.

Even though I cannot predict the results of some of the studies I suggested in Section 2, I predict that I will be happy to go along with the folk, as long as they are consistent. Part of my reason for that prediction is my confidence that (setting aside so-called negative actions again) if there is a folk concept of intentional action, it has at its core intentional action of the sort that has been my primary concern, and the primary concern of many philosophers of action, all along – the kind I had in mind in my efforts to understand how agents produce their intentional actions. If the folk concept of intentional action includes that kind of intentional action at its core, the concept’s ranging more broadly is no special cause for worry for philosophers whose project is to explain how intentional action is produced.

Acknowledgments

For comments on a draft of this article, I am grateful to Joshua Knobe, Gideon Yaffe, and two anonymous referees.

Notes

¹ I have reviewed some of it elsewhere (Mele 1992b, 1997).

² For roughly this understanding of folk concepts, see Jackson 1998, ch. 2.

³ If hitting the bull's-eye seems to be a more difficult task than killing the woman, that may account for part of the difference in the respondents' judgments. Knobe also asked about variants of these cases in which "Jake is an expert marksman. His hands are steady. The gun is aimed perfectly." 95% of the respondents said that he intentionally killed his aunt and 75% said that he intentionally hit the bull's-eye. This difference in responses might be partially accounted for by differences in respondents' assumptions about difficulty. In an ideal test, it would be clear that the moral and nonmoral tasks are equally difficult. For a pair of cases of this kind, see Mele & Sverdlik 1996, pp. 281-82.

⁴ In the passages I quoted from Harman 1976, pp. 433-44, both reasons and "what ought to be" reasons are mentioned. Discussion of the nature of reasons and of competing views about that is beyond the scope of this article.

⁵ Knobe reports that in the Jake study "91% of subjects said that [he] was trying" (2003, p. 000). In personal correspondence he reported the breakdown: 95% said that he was trying to hit the bull's-eye and 88% said that he was trying to kill his aunt.

⁶ This thesis resembles Bratman's "Single Phenomenon View," the thesis that "to *A* intentionally I must intend to do something" (1987, p. 113). Bratman advances it as an alternative to the Simple View.

⁷ On negative actions, see Mele 2003, pp. 146-54. This is a topic that cries out for studies like Knobe's. If it were to turn out that the majority of lay folk describe certain not-doings as intentional while denying them the status of actions, that would be interesting.

⁸ To forestall confusion, I emphasize that as I understand the notion of a property's lying at the core of the folk concept of intentional action, to say that a property *P* lies at the core of the concept is to say that (setting aside negative actions) no possible case of action that instantiates the concept lacks *P*.

⁹ I assume that the great majority of people who say that Jake intentionally killed his aunt would also say that he intentionally shot his aunt. If the assumption is false, we obviously have a puzzle on our hands.

¹⁰ Robert Audi argues that there are cases in which an agent acts intentionally without intending to do anything (1986, pp. 27-28). For criticism, see Mele 1992a, pp. 186-87. Incidentally, Audi's thesis that *S* intends at *t* to *A* only if "at *t* . . . *S* believes that he will (or that he probably will) *A*" (p. 18) plays a role in his argument. Whether that thesis is consistent with the folk concept of intention is worth testing.

¹¹ For discussion of a seemingly problematic case, see Mele 1992a, pp. 149-50.

References

Adams, F. (1986). "Intention and intentional action: the simple view." *Mind and Language* 1: 281-301.

-
- Anscombe, G. E. M. (1963). *Intention*. 2nd ed. Ithaca, NY: Cornell University Press.
- Audi, R. (1986). "Intending, intentional action, and desire." In J. Marks, ed. *The ways of desire*. Chicago: Precedent.
- Bratman, M. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- _____. (1984). "Two faces of intention." *Philosophical Review* 93: 375-405.
- Butler, R. (1978). "Report on *Analysis* problem no. 16." *Analysis* 38: 113-14.
- Chisholm, R. (1966). "Freedom and action." in K. Lehrer, ed. *Freedom and determinism*. New York: Random House.
- Davidson, D. (1980). *Essays on actions and events*. Oxford: Clarendon Press.
- Ginet, C. (1990). *On action*. Cambridge: Cambridge University Press.
- Goldman, A. (1970). *A theory of human action*. Englewood Cliffs, NJ: Prentice-Hall.
- Harman, G. (1976). "Practical reasoning." *Review of Metaphysics* 79: 431-63.
- Jackson, F. (1998). *From metaphysics to ethics: a defence of conceptual analysis*. Oxford: Clarendon.
- Knobe, J. (2003). "[Intentional action in folk psychology: an experimental investigation.](#)" *Philosophical Psychology*.
- _____. (n.d.-a). "[Intentional action and side-effects in ordinary language.](#)" *Analysis*, forthcoming.
- _____. (n.d.-b). "Comparing 'intentionally' and 'intention'." Unpublished ms.
- McCann, H. (1991). "Settled objectives and rational constraints." *American Philosophical Quarterly* 28: 24-36.
- _____. (1986). "Rationality and the range of intention." *Midwest Studies in Philosophy* 10: 191-211.

-
- Mele, A. 2003. *Motivation and agency*. New York: Oxford University Press.
- _____. (2001). "Acting intentionally: probing folk notions," in B. Malle, L. Moses, and D. Baldwin, eds. *Intentions and intentionality: foundations of social cognition*. Cambridge, MA: MIT Press.
- _____. (1997). "Introduction." In A. Mele, ed. *The philosophy of action*. Oxford: Oxford University Press.
- _____. (1992a). *Springs of action*. New York: Oxford University Press.
- _____. (1992b). "Recent work on intentional action." *American Philosophical Quarterly* 29: 199-217.
- Mele, A. & S. Sverdlik. (1996). "Intention, intentional action, and moral responsibility." *Philosophical Studies* 82: 265-87.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
- Scanlon, T. M. (1998). *What we owe to each other*. Cambridge, MA: Harvard University Press.
- Strawson, G. (1994). "The impossibility of moral responsibility." *Philosophical Studies* 75: 5-24.
- Thalberg, I. (1977). *Perception, emotion, and action*. New Haven, CT: Yale University Press.
- Thomson, J. (1977). *Acts and other events*. Ithaca, NY: Cornell University Press.