Can One Act for a Reason without Acting Intentionally?

Joshua Knobe and Sean D. Kelly

Princeton University

Since the important work of Elizabeth Anscombe (1957), philosophers have been almost unanimous in accepting the claim that every behavior performed for a reason is performed intentionally. We develop and discuss here an apparent counterexample to this claim – a case in which people are inclined to say that an agent has performed a behavior for a reason, but are not inclined to say that he has performed that behavior intentionally. Absent a convincing argument for the view that people's intuitions are misguided in this case, one must conclude that Anscombe's claim is incorrect.

I

It is widely agreed that the concept of acting intentionally is closely linked in some way to the concept of acting for a reason, but there has been a certain amount of disagreement about the precise nature of this link. We therefore begin by distinguishing two related theses:

> (*Davidson's Thesis*) Every behavior that is performed intentionally is performed for a reason.
>
> (*Anscombe's Thesis*) Every behavior that is performed for a reason is performed intentionally.

Taken together, Davidson's Thesis and Anscombe's Thesis entail the biconditional conclusion that a behavior is performed intentionally if and only if it is performed for a reason. This is a conclusion of great theoretical power and simplicity, and it has probably been discussed more widely than either of the two individual theses taken separately. Still, it is important to remember

that the two theses are logically independent. Arguments against one of the two theses will not usually count as arguments against the other.

Although Davidson's Thesis has received vigorous support from a number of philosophers of action (e.g., Davidson 1980; Goldman 1970; Mele 1992), it has also occasionally been called into question. Hursthouse (1991), for example, claims that spontaneous expressions of emotion are not performed for any reason but still count as intentional actions. Consider the agent who gets very angry and ends up punching the wall. It seems clear that such an agent has punched the wall intentionally, but on Hursthouse's view at least, he or she has not punched the wall for any reason.

Here, however, our primary concern is not with Davidson's Thesis but with Anscombe's Thesis. This latter thesis has not been particularly controversial. It is upheld by Goldman (1970), Malle, Knobe, O'Laughlin, Pearce and Nelson (2000), Mele (1992) and, of course, Anscombe (1957). We know of no arguments against it.[1]

A key source of evidence for Anscombe's Thesis is our ordinary practice of giving reason explanations. Specifically, it is thought that people do not ordinarily accept reason explanations for behaviors that they do not regard as intentional. If a speaker thinks that Jane bumped into

[1] However, for an argument in the general area see Mele, "Acting for Reasons and Acting Intentionally" (APQ, 1992). Mele claims that in some lucky achievement cases – cases of simple luck and cases of causal deviance – people may have middling opinions on average about whether the action was done intentionally but be strongly inclined to see the action as done for a reason. For example, in a deviance case subjects may have middling opinions about whether the shooter hit the bull's-eye intentionally but be strongly inclined to say that he hit it in order to win the prize. It would be interesting to test such cases.

Bob unintentionally, that speaker will not normally find it acceptable to use a reason explanation like 'Jane bumped into Bob in order to get his attention.' As long as Jane's bumping into Bob is regarded as an unintentional behavior, it will be felt that Jane did not bump into Bob for any reason at all and hence that it would be wrong to say that Jane bumped into Bob 'in order to' fulfill any kind of purpose. Similarly with many other types of unintentional behavior. If Jane unintentionally shivers or trips or starts crying, it will be felt that she is not performing these behaviors for any reason and that it would therefore be wrong to explain her behavior using phrases like 'in order to' that are ordinarily reserved for reason explanations.

But, of course, accidental behaviors and uncontrollable reflexes are only the simplest forms of unintentional behavior. The true test of Anscombe's Thesis is its ability to handle more complex cases — cases in which a behavior would normally be considered unintentional even though it was foreseen and produced by a deliberate choice.

II

We turn, therefore, to the difficult problem of *side effects*. An outcome can be considered a side effect if and only if

(1) an agent chooses to perform a behavior, foreseeing that she will thereby (probably) bring about the outcome, but

(2) the agent does not actually perform the behavior for the purpose of bringing about the outcome.

Thus, suppose that the chairman of the board of a company decides to implement a particular program. He knows that this program will lead to some outcome *x*. But he doesn't care at all

about outcome *x*; he has only chosen to implement the program because he thinks that he can thereby increase profits. Here we can say that outcome *x* is a 'side effect' of the chairman's behavior.

A question now arises as to whether side effects can ever be brought about intentionally. Here, opinions differ. Adams (1986) says that side effects are never brought about intentionally; Ginet (1990) says that side effects are always brought about intentionally; Harman (1976), Bratman (1984; 1987) and Knobe (2003) say that some side effects are brought about intentionally and others are brought about unintentionally.

As we shall see, people's ordinary use of the word 'intentionally' follows this third view. People are willing to apply the word to some side effects but not to others. Thus, there can be no single general answer as to whether or not people think that the chairman intentionally brought about 'some outcome *x*.' People's answers will depend in a crucial way on what outcome *x* happens to be.

This fact about people's use of 'intentionally' gives us a valuable opportunity to put Anscombe's Thesis to the test. We can check to see whether people's use of reason explanations fits with their use of 'intentionally' in the way Anscombe's Thesis seems to suggest. Of course, Anscombe's Thesis could potentially be true even if it does not comport well with ordinary language, but it seems clear that ordinary language provides at least prima facie evidence regarding difficult questions like this one. Suppose we found that people were sometimes willing to accept reason explanations when they felt that a side effect had been brought about

intentionally but that they were never willing to accept reason explanations when they felt that a side effect had been brought about unintentionally. Surely, this finding would provide a kind of evidence for Anscombe's Thesis — not conclusive evidence, of course, but evidence nonetheless. And now consider another possible result. Suppose we found that people sometimes accepted reason explanations for side effects that they regarded as unintentional. We would then have a very strong argument against Anscombe's Thesis. Indeed, the only way to defend Anscombe's Thesis against the argument would be to show either that ordinary English speakers were mistaken in their use of the word 'intentionally' or that they were mistaken in their use of reason explanations. Although it might be possible to show that ordinary speakers are mistaken in one of these ways, we take it that there is a very strong presumption in favor of the view that ordinary speakers are speaking correctly.

With these issues in mind, let us consider two specific cases in which an agent brings about a side effect. First, a case in which the side effect is environmental 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.

And now consider what happens when we modify this case by replacing 'harm' with 'help':

> 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 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.

It has been shown experimentally that these two cases yield two radically different patterns of intuition about whether or not the agent acted intentionally (Knobe 2003). Most people who receive the first case say that the chairman *intentionally* harmed the environment, but most people who receive the second case say that the chairman *unintentionally* helped the environment. Presumably, the key difference between these two cases has something to do with the normative status of the two outcomes. Because people see environmental harm as a bad thing and environmental help as a good thing, they end up concluding that the former was brought about intentionally, the latter unintentionally.

There is, as yet, no consensus on the question as to why normative considerations affect people's responses to this pair of cases. Some researchers say that normative considerations actually play a role in people's concept of intentional action (Knobe forthcoming; Mele 2003; Nadelhoffer forthcoming); others say that people are being unduly influenced by feelings of blame (Malle forthcoming; Malle & Nelson 2003); and still others say that the effect can be explained in terms of conversational pragmatics (Adams & Steadman forthcoming A;

forthcoming B).[2] We will make no attempt to resolve this controversy here. Instead, we want to focus on the question as to whether people's intuitions about intentional action — however these intuitions might have arisen — end up fitting with their intuitions about reason explanations in the way that Anscombe's Thesis seems to suggest.

Assuming that people's intuitions accord with Anscombe's Thesis, it seems that people may feel that the chairman harmed the environment for a reason but that they definitely shouldn't feel that the chairman helped the environment for a reason. In fact, people's intuitions fit this pattern perfectly. Experimental results show that people think it sounds right to say 'The chairman *harmed* the environment in order to increase profits' but that people do not think it sounds right to say 'The chairman *helped* the environment in order to increase profits' (Knobe, forthcoming).

This is truly an impressive victory for Anscombe's Thesis. Indeed, it might be thought that Anscombe's Thesis provides the only natural way of explaining the asymmetry in people's use of 'in order to' in these two cases. The natural explanation is this. In both cases it is clear that the chairman's desire to increase profits led ultimately to the side effect in question. But since people think that he didn't help the environment intentionally, and since Anscombe's Thesis says that every behavior done for a reason is done intentionally, it seems perfectly reasonable for people to conclude that the chairman didn't help the environment for a reason.

[2] Note that although some of these theories assert that people incorrectly classify immoral behaviors as *intentional*, none of them asserts that people incorrectly classify morally good behaviors as *unintentional*. The counterexample we propose in section III involves the classification of a certain behavior as unintentional. Thus there is nothing in any of these theories to suggest that that classification is incorrect.

III

Although the cases we have just discussed seem to be evidence for Anscombe's Thesis, there is even stronger evidence that the thesis is false. One can construct cases in which subjects judge that an agent has performed an action for a reason even though they also judge that he did not perform it intentionally. In the cases of the chairman and the environment, as we have seen, judgments about acting for a reason and acting intentionally line up nicely. But this is only because in these cases there is no radical divergence between two different kinds of normative evaluation that the subject can make. When we devise examples that pull these normative evaluations apart, we get robust evidence against Anscombe's Thesis.

The two normative evaluations in question are the evaluations of the moral goodness of the side-effect as understood by the subject, on the one hand, and as understood by the agent, on the other. In the cases discussed in section III these evaluations line up fairly well; or at least they do not diverge radically. To see this, consider the actions of helping and harming the environment from the point of view of normal subjects and of the chairman in the two scenarios.

We may assume that in general subjects think that helping the environment is better than harming it. This is a fairly prevalent view in our society, and it seems reasonable to think that most subjects have something like this as a background assumption. (In future experiments, of

course, this is the kind of thing that might be tested for explicitly.) Likewise, in the two chairman scenarios there is no strong reason to think that the chairman evaluates these actions any differently than the subjects do. In any case, he certainly does not make the opposite evaluation: there is no reason at all to think that he believes that harming the environment is better than helping it. He is, of course, at the very least callous in the case in which he goes ahead and harms the environment for profit. But he is not malicious; he is not represented as taking any joy in harming the environment, and indeed he says that he doesn't care one way or the other about it.

But it is possible to devise scenarios in which the agent's evaluation of the moral worth of a side-effect is represented as explicitly contravening the subject's (assumed) evaluation of the very same act. This will work best if the agent is represented as someone with moral values that are diametrically opposed to our own. Consider, for instance, the following scenario.

> A terrorist discovers that someone has planted a bomb in a nightclub. There are lots of Americans in the nightclub who will be injured or killed if the bomb goes off. The terrorist says to himself, "Whoever planted that bomb in the nightclub did a good thing. Americans are evil! The world will be a better place when more of them are injured or dead."
>
> Later, the terrorist discovers that his only son, whom he loves dearly, is in the nightclub as well. If the bomb goes off, his son will certainly be injured or killed. The terrorist then says to himself, "The only way I can save my son is to

defuse the bomb. But if I defuse the bomb, I'll be saving those evil Americans as well... What should I do?"

After carefully considering the matter, he thinks to himself, "I know it is wrong to save Americans, but I can't rescue my son without saving those Americans as well. I guess I'll just have to defuse the bomb."

He defuses the bomb, and all of the Americans are saved.

In this scenario, the side-effect is that the Americans are saved. Unlike in the chairman cases, however, it seems here that the agent's evaluation of the moral worth of this side-effect explicitly contravenes the evaluation of normal American subjects. For although we can assume that a typical sample of American subjects will judge that it is a good thing to save the lives of innocent Americans, the terrorist is represented as believing that this side-effect is morally bad.

Now we can ask two distinct questions about this side-effect. First, does the terrorist *intentionally* save the Americans? And second, does the terrorist save the Americans *for a reason* (namely in order to rescue his son)? Among ordinary speakers of the English language the answers to these questions diverge dramatically.

To demonstrate this we conducted a simple experiment. We asked 62 subjects in two New York City parks to read the above scenario and fill out a brief questionnaire. We randomly assigned subjects either to the 'reason' condition or to the 'intentionally' condition. We asked

subjects in each condition to evaluate a certain sentence. In the 'reason' condition subjects evaluated the sentence "The terrorist saved the Americans in order to rescue his son." In the 'intentionally' condition subjects evaluated the sentence "The terrorist saved the Americans intentionally." Evaluations consisted in rating the sentence on a scale from –3 ('sounds wrong') to +3 ('sounds right'), with the 0 point marked 'in between.' Subjects in both conditions were also asked to evaluate two sentences concerning the normative status of the terrorist's actions. These sentences were (1) "The terrorist deserves praise for saving the Americans" and (2) "The terrorist deserves blame for violating his own moral code."

Experimental results for this case call Anscombe's Thesis into question. Although subjects tended strongly to judge that the terrorist saved the Americans *for a reason* (namely, in order to rescue his son), they also tended to judge that he did not save the Americans *intentionally*. The average rating for the sentence "The terrorist saved the Americans in order to rescue his son" was 1.8. The average rating for the sentence "The terrorist saved the Americans intentionally" was –0.5. This difference was highly statistically significant, t (62)=4.13, p<.001.

In the table below we have combined all negative ratings under the heading 'sounds wrong' and all positive ratings under the heading 'sounds right'.

	'intentionally'	'reason'
'sounds right'	33.1%	78.1%
'in between'	16.7%	6.3%

'sounds wrong'	50%	15.6%

The divergence of responses in the two conditions is striking.[3] They seem to provide an example in which it is natural to say that an agent has acted for a reason, but natural also to say that he has not acted intentionally. They provide robust evidence, in other words, against Anscombe's Thesis. In the following section we will analyze these results.

IV

We have seen that normative judgments influence people's intuitions about the application of two distinct action-theoretic concepts: the concept of acting *intentionally* and the concept of acting *for a reason*. But we have also seen that a subject's intuition about how to apply each of these concepts is influenced by a different normative judgment:

[3] Interestingly, the results for the 'intentionally' condition differed between the two parks in which the experiment was conducted. The two parks were Central Park (the part behind the Metropolitan Museum of Art) and Washington Square Park (in Greenwich Village). In the Central Park sample there were 28 subjects, and the average rating for the sentence in the 'intentionally' condition was -1.7; in the Washington Square Park sample there were 34 subjects, and the average rating for the sentence in the 'intentionally' condition was 0.4. To test for a difference between parks we ran a 2 x 2 ANOVA test with condition ('intentionally' vs. 'reason') and sample ('Washington Square Park' vs. 'Central Park') as independent variables. There was a highly significant main effect for condition, $F(1, 61) = 18.5$, $p < .001$. But the main effect for sample fell short of significance $F(1, 61) = 3.1$, $p = .08$, as did the interaction between condition and sample $F(1, 61) = 3.3$, $p = .07$.

1) Intuitions about whether a side-effect was brought about intentionally are influenced by the *subject's* judgment about its normative status.

2) Intuitions about whether a side-effect was brought about for a reason are influenced by what the subject believes to be the *agent's* judgment about its normative status.

Although different normative judgments are relevant to subjects' intuitions about acting intentionally and acting for a reason, there is a general principle about how the normative status of a side-effect influences the application of action-theoretic concepts. This principle is general in the sense that it applies to both of our action-theoretic concepts – acting intentionally and acting for a reason. Once we have determined which normative judgment is relevant, the principle predicts whether or not a subject will attribute the corresponding action-theoretic concept to a given side-effect. The general principle in question is:

Principle of Un-charity: A subject is more willing to attribute the action-theoretic concept to a side-effect if the side-effect violates the relevant norm.

When combined with (1) and (2) above, the Principle of Un-charity yields definite predictions about whether or not a subject will attribute each of the action-theoretic concepts to a given side-effect: if the side-effect violates the *subject's* norms for behavior, he will judge it to have been brought about *intentionally*; if it violates the *agent's* norms for behavior, he will judge it to have been brought about *for a reason*.

In determining whether a side-effect was brought about intentionally or for a reason the subject seems to go through a two-step process. First, he chooses the normative evaluation of the side-effect that is most relevant: his own or the agent's. Next, he uses that normative evaluation in accordance with the Principle of Un-charity: if the side-effect violates the relevant norm, he is more willing to attribute the action-theoretic concept in question. Let us consider each of these steps in turn.

First, the choice of the relevant norm. Suppose the subject is considering whether a side-effect was brought about intentionally. Why should the *subject's* normative evaluation be relevant in this case? There are many possible explanations here, but one plausible account goes like this. When a subject is trying to determine whether a behavior was performed intentionally, it is often because he is trying to decide whether the agent deserves praise or blame for that behavior. In determining whether a behavior is praiseworthy or blameworthy, moreover, the *subject's* normative evaluation of the behavior is the most salient one. (Would you really praise an agent for performing an action that you yourself didn't believe was good?) For this reason, the subject's normative evaluation of the behavior is the most salient one in determining whether the behavior was performed intentionally.

Suppose, now, that the subject is considering whether a side-effect was brought about for a reason. We have seen in this case that the subject considers the *agent's* norm to be the relevant one. Why should this be so? Again, there are many possible explanations. Consider the following plausible account. When a subject is trying to determine whether a behavior was performed for a reason, it is natural for her to look at the behavior from the agent's point of view.

The question whether the behavior was done for a reason is understood as a question about the *agent's* reasons for performing it. Therefore, the agent's normative evaluation of the behavior is the most salient one in determining whether the behavior was performed for a reason.

Let us turn now to the Principle of Un-charity. As we have seen, this principle states that a subject is more willing to attribute certain action-theoretic concepts to a side-effect if that side-effect violates a relevant norm. The principle seems to apply both to the concept of intentional action and to the concept of acting for a reason. Why should it apply to both of these concepts?

First, the case of intentional action. The Principle of Un-charity implies an asymmetry here: side-effects that the subject regards as bad are judged to have been brought about intentionally while side-effects that the subject regards as good are judged to have been brought about unintentionally. One plausible explanation for this asymmetry in people's judgments about intentional action ties it to a prior asymmetry in judgments of praiseworthiness and blameworthiness. Subjects are generally more willing to assign blame for side-effects that they regard as bad than they are to assign praise for side-effects that they regard as good. This may be a purely contingent fact about our assignments of praise and blame; but it is a conspicuous and notable fact nonetheless. Recall the scenarios from sections II and III. Subjects are much more willing to blame the chairman when he harms the environment (a bad side-effect) than they are to praise him when he manages to help it (a good side-effect). Similarly, subjects do not regard the terrorist as worthy of praise, even though his behavior brings about the good side-effect of

saving the Americans.[4] As we have seen already, there are a variety of different theories that attempt to explain the interaction between judgments of praise and blame on the one hand, and judgments of intentional action on the other. But all of these theories admit that there is an interaction between these judgments, and on all of these theories the asymmetry that we have noted between judgments of praise and blame can be used to explain the asymmetry in judgments about intentional action.

Finally, the case of acting for a reason. We need to explain why subjects are more willing to say that a side-effect was brought about for a reason when the side-effect violates the agent's norms. We have seen that the question whether the side-effect was brought about for a reason amounts, in this case, to a question about the agent's own reasons for bringing it about. The agent will feel especially compelled to provide such reasons when his behavior brings about a side-effect that he considers bad. Reason explanations, in cases like this, can serve to justify the side-effect that the agent's behavior brings about. A subject who is trying to determine whether a behavior was performed for a reason will naturally think of the situation from the agent's point of view. She will therefore be more willing to provide reason explanations in cases in which the agent needs to justify his action.

---

[4] Indeed this tendency was apparent within the experiment itself. Subjects were asked to evaluate the sentence "The terrorist deserves praise for saving the Americans." On a scale from –3 ("sounds wrong") to +3 ("sounds right"), they gave an average rating of –1.6.

One additional fact is worth reporting. Subjects tended not to agree with the statement "The terrorist deserves blame for violating his own moral code." On a scale of –3 to +3 they gave this statement an average rating of –1.7. As will become clear, we propose that the subject thinks of the terrorist as having had an internal reason for performing what he (the terrorist) understands to be a morally bad act. In other words, he is understood as having an excuse. Whether in addition he deserves blame for having performed (what he understands to be) a morally bad act depends upon whether that excuse was a *good* one. Our view takes no stand on this question.

V

The experiments reported here suggest that subjects are sometimes willing to say that an agent has performed a behavior for a reason even though they are not also willing to say that he performed the behavior intentionally. We believe that these results count as evidence against Anscombe's Thesis – the thesis that every behavior performed for a reason is performed intentionally. If subjects are systematically willing to use the expression "in order to" in cases in which they are not also willing to use the expression "intentionally," then subjects' use of these terms seems to be evidence against Anscombe's Thesis.

How might an objector reply to our claims? If our experiments have been conducted properly, they do seem to establish an interesting fact about normal linguistic use. It remains open to an objector, however, to insist that facts about normal linguistic use are no sure guide to the truth of Anscombe's Thesis. The thesis that every behavior performed for a reason is performed intentionally, the objector might say, is not a thesis about how people use the relevant expressions but a thesis about what is true of our behaviors. Take, by analogy, the case of arthritis. If we want to understand the true nature of this disease, it would be folly to proceed by investigating the ordinary use of the term "arthritis." So too, the objector will say, facts about the ordinary use of expressions like "in order to" are simply irrelevant to questions about the true nature of human action.[5]

[5] See Weatherson (2003), who argues that people's intuitions about the application of certain predicates can systematically diverge from the conditions under which those predicates are properly applied**.**

An objection like this does present a legitimate challenge to our view. In particular, it shows that our argument against Anscombe's Thesis is not *conclusive*. Still, we believe that the evidence we have presented here serves at the very least to shift the burden of proof. For now a person who wants to hold on to Anscombe's Thesis needs to tell a convincing story about why normal linguistic use diverges from it so systematically. Such a story might be forthcoming, but in its absence the evidence from linguistic use seems hard to ignore.

Bibliography

Adams, Fred (1986). Intention and Intentional Action: The Simple View. *Mind and Language*, 1, 281-301.

Adams, Fred and Annie Steadman (forthcoming A). Intentional Action in Ordinary Language: Core Concept or Pragmatic Understanding? *Analysis*.

Adams, Fred and Annie Steadman (forthcoming B). Intentional Actions and Moral Considerations: Still Pragmatic. *Analysis*.

Anscombe, G. E. M. (1957). *Intention*. Ithaca: Cornell University Press.

Bratman, Michael (1984). Two Faces of Intention. *Philosophical Review*, 93, 375-405.

Bratman, Michael (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard

University Press.

Davidson, Donald (1980). Actions, Reasons, and Causes. Reprinted in *Essays on Actions and Events*. Oxford: Clarendon Press.

Ginet, Carl. (1990). *On Action*. Cambridge: Cambridge University Press.

Goldman, Alvin (1970). *A Theory of Human Action*. Englewood Cliffs: Prentice Hall.

Harman, Gilbert (1976). Practical Reasoning. *Review of Metaphysics*, *29*, 431-463.

Hursthouse, Rosalind (1991). Arational Actions. *Journal of Philosophy*, LXXXVIII, 2, 57-68.

Knobe, Joshua (forthcoming). Intention, Intentional Action and Moral Considerations. *Analysis*.

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

Malle, Bertram (forthcoming). *How People Explain Behavior*. Cambridge, MA: MIT Press.

Malle, Bertram F., & Sarah E. Nelson (2003). Judging *Mens Rea*: The Tension Between Folk Concepts and Legal Concepts of Intentionality. *Behavioral Sciences and the Law*, 21, 563-580.

Malle, Bertram F., Joshua Knobe, Matthew O'Laughlin , Gale Pearce, & Sarah Nelson (2000). Conceptual Structure and Social Functions of Behavior Explanations. *Journal of Personality and Social Psychology*, 79, 309-326.

Mele, Alfred (1992). Acting for Reasons and Acting Intentionally. *Pacific Philosophical Quarterly*, 73, 355-374.

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

Nadelhoffer, Thomas (forthcoming). The Butler Problem Revisited. *Analysis*.

Weatherson, Brian 2003. What Good are Counterexamples. *Philosophical Studies*, 115 (2003): 1-31.