Intentional Action and the Unintentional Fallacy

Ryan Wasserman
Western Washington University

1. Introduction

Anscombe begins her book *Intention* by drawing a distinction between *intentions* and *intentional actions*, noting that “we may be tempted to think that only actions done with... intentions ought to be called intentional.”¹ This is an early statement of what Michael Bratman calls “the Simple View”:

*The Simple View*: S φ-ed intentionally only if S intended to φ.²

Anscombe and Bratman both reject the Simple View, but many philosophers have thought that it holds the key to understanding intentional action.³

The Simple View raises a complex question: If intentional actions require intentions, then what do intentions require? In other words, what are the necessary conditions for intending to φ? For simplicity, I will assume a standard belief-desire model according to which S intends to φ only if S desires to φ and believes that he will do so.⁴

Much of the recent work in action theory can be organized around a set of objections facing the Simple View and other intention-based accounts of intentional action.⁵ My aim is to present the three most common objections

---

¹ Anscombe (1957, p.1).
³ Chisholm (1964), for example, suggested that an agent performs an action intentionally just in case he performs the action in the way he intended.
⁴ See, e.g., Davis (1984). This assumption is controversial, but it will make my project harder, rather than easier. The assumption is therefore unobjectionable, at least for the purposes of this paper.
⁵ Mele (1992) provides a helpful survey.
to the Simple View (sections 2-4) and to argue that all three objections commit a common fallacy (sections 5-6). I will then draw some more general conclusions about the connection between intentional action and moral responsibility (section 7).

2. The Problem of Side-Effects

Gilbert Harman famously describes a case in which a sniper shoots an opposing soldier, knowing that this will alert his enemies to his position. “He does this intentionally,” Harman suggests, “thinking that the gain is worth the possible cost.” As Harman notes, it is natural to describe the sniper as having alerted the enemy intentionally. There is also some pressure to deny that he intended for this to happen, since he lacked the requisite desire. Putting the two pieces together, we have an apparent counterexample to the Simple View, for we have an intentional action that was not intended by the agent. Similar examples are discussed by many philosophers including Carl Ginet, who concludes that:

In general, side effects of an intended action that the agent expects (in the sense that she is aware at the time of choosing the action that she will or might thereby bring them about) need not therefore be effects the agent intends to bring about, but they are therefore effects she brings about intentionally.

More recently, several authors have noted that moral considerations can magnify the problem of side-effects. Much of the recent literature has focused on the following pair of cases, devised by Joshua Knobe:

_Harm Vignette:_ 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.

---

8 See Knobe (2003) and McCann (2005). Harman stresses the same point in his original article.
Help Vignette: The vice-president of a company went to the chairman of the board and said, ‘we are thinking of starting a new program. It will help us increase profits, but it will also help the environment.’ The chairman of the board answered, ‘I don’t care at all about helping the environment. I just want to make as much profit as I can. Let’s start the new program.’ They started the new program. Sure enough, the environment was helped.⁹

The first case is compelling—Knobe reports that 82% of respondents describe the chairman as having intentionally harmed the environment. But given our current assumptions, the chairman did not intend to harm the environment since he did not desire this result.¹⁰ Interestingly, the Help Vignette elicits very different reactions, with only 23% of respondents describing the chairman as having intentionally helped the environment.¹¹ An adequate defense of the Simple View must explain our mistaken judgment in the first case, but it should also account for the striking asymmetry uncovered by Knobe.

3. The Problem of Luck

The problem of side-effects focuses on consequences that are unwanted; the problem of luck concerns ones that are unlikely. Consider the following example.¹² Time is running out in the basketball game. Ned has the ball at half-court and, knowing that the only chance for victory is a last second shot, heaves the ball toward the hoop. Miraculously, the ball goes in. Did Ned hit the shot intentionally? Opinions differ,¹³ but those tempted to answer affirmatively have a second reason to doubt the Simple View. Ned did not have an outright belief that he would hit the shot, since he recognized the

---

¹⁰ Actually, matters are not entirely clear at this point. McCann (2005) reports that a significant number of respondents (42%) credit the vice-president with having intended to harm the environment in the harm vignette.
¹¹ The same kind of asymmetry is noted by Harman in connection with his original case. The sniper knows that shooting the soldier will both alert the enemy to his position and heat the barrel of this gun. Both side-effects are foreseen; neither is desired. Yet many of us are inclined to say that only the first is brought about intentionally.
¹² A similar case is discussed by Harman (1976, p.433).
improbability of success. Thus, given our current assumptions, this is another example of an intentional action that was not intended by the agent.

As in the case of side-effects, moral considerations help to dramatize the problem of luck. Consider, for example, the following puzzle from Ronald Butler:

If Brown in an ordinary game of dice hopes to throw a six and does so, we do not say that he threw the six intentionally. On the other hand if Brown puts one live cartridge into a six-chambered revolver, spins the chamber as he aims it at Smith and pulls the trigger hoping to kill Smith, we would say if he succeeded that he had killed Smith intentionally. How can this be so, since in both cases the probability of the desired result is the same?\(^{14}\)

Once again, there are two challenges for the defender of the Simple View. First, one must explain our mistaken judgment that Brown shot Smith intentionally. Second, one should explain the interesting asymmetry identified by Butler.

4. The Problem of Consistency

Michael Bratman presents a third and final objection to the Simple View that begins with an intuitive constraint on intending:

*The Strong Consistency Requirement:* \(S\) is rational only if \(S\)'s intentions can be combined into an overall plan that is consistent with \(S\)'s beliefs.\(^{15}\)

Bratman next imagines an unusual example in which he ambidextrously plays two videogames at once. In each hand he holds a controller, with each

\(^{14}\) Butler (1976, p.113). Harman makes the same point by replacing the basketball player in the earlier example with a sniper and the basketball hoop with an enemy soldier. The shot in this example is taken from a great distance and the sniper doubts its success. Harman notes that if the shot is successful, we will credit the sniper as having killed intentionally despite his lack of faith.

\(^{15}\) Bratman (1984, p.380). If intending requires belief, as I assume throughout, we have a straightforward explanation of the normative force of this constraint, since the rational agent cannot believe (a) that he will \(\phi\), (b) that he will \(\psi\) and (c) that \(\phi\)-ing and \(\psi\)-ing are incompatible. Note also that SCR helps to explain the difference.
controller linked to a separate television screen. There is a target on each screen and in order for Bratman to win one of the games he must use the corresponding controller to guide a “missile” to the target. The twist is that the games are linked, so that it is impossible to hit both targets (if both missiles are going to strike, the game automatically shuts down). Suppose that this is known to the agent. Suppose also that the targets are difficult to hit and that the odds of success are therefore maximized by playing both games at once (that is, by trying to hit Target 1 and, at the same time, trying to hit Target 2). Suppose that Bratman does just this and manages to hit Target 1 in exactly the way that he envisioned. In that case, argues Bratman, he should be credited with having hit Target 1 intentionally.

The problem for the Simple View is clear. If Bratman hit Target 1 intentionally, he intended to hit Target 1. But if he intended to hit Target 1, he must have intended to hit Target 2 as well, since his attitudes toward the two targets were the same. The Strong Consistency Requirement then tells us that Bratman is irrational, for this pair of intentions cannot be combined into an overall plan consistent with his beliefs about the setup of the videogame system. But Bratman pleads innocence on the grounds that “the strategy of giving each game a try seems perfectly reasonable.” The Simple View is thus subject to a reductio; Bratman hits Target 1 intentionally, without having had the intention to hit that target.

Note that Bratman’s objection can be sharpened by introducing moral considerations. Replace the targets with people and the videogame missiles with genuine missiles. If Bratman successfully strikes Person 1, it will be difficult to deny that he does so intentionally.

5. The Unintentional Fallacy

Taken together, the problems of side-effects, luck, and consistency present a serious challenge to the Simple View. Fortunately, the challenge can be met since all three objections commit a common error that I call the Unintentional Fallacy.

It is a familiar fact that ‘not’ is typically interpreted as taking narrow scope when interacting with propositional attitude verbs. Consider:

---

17 Note that, unlike the previous two problems, Bratman’s objection does not make any direct appeal to the belief-desire model of intention.
18 The phenomenon is not limited to propositional attitude verbs, of course. If one denies that Sally should take an umbrella, one will communicate that Sally shouldn’t take the umbrella.
(1) Al thinks it will rain.
(2) Al does not think it will rain.
(3) Al thinks it won’t rain.

The denial of (1) implies (2), which is consistent with agnosticism; but it communicates (3), which requires the contrary belief.

The same pattern holds for many adverbs. Consider:

(4) Bill went to work happily.
(5) Bill did not go to work happily.
(6) Bill went to work unhappily.

The denial of (4) implies (5), which is consistent with ambivalence; but it communicates (6), which requires the contrary attitude.

Now consider the case of intentional action:

(7) Bill broke the window intentionally.
(8) Bill did not break the window intentionally.
(9) Bill broke the window unintentionally.

The denial of (7) implies (8), which is consistent with both trying and foresight; but it communicates (9), which requires the damages to be accidental.

In all of these cases, we unconsciously treat contraries like “happy” and “unhappy” like contradictories and equate the denial of one with the affirmation of the other. This kind of inference is perfectly natural. It might even be inductively strong. But the point to stress is that it is not logically sound. The distinguishing feature of contrary statements like (7) and (9) is that they cannot both be true, but they can both be false. The middle is not excluded, so one cannot deduce (9) from the denial of (7). To do so is to commit the Unintentional Fallacy.¹⁹

This phenomenon is familiar, but it raises many interesting questions. Clearly, many contexts license the assumption \( A \text{ or } B \), where \( A \) and \( B \) are contraries like (1) and (3). But when, exactly, is such an assumption warranted? Is the assumption a semantic presupposition or something

---

pragmatic? And how did this pattern of inference get started in the first place? Fortunately, we need not answer these questions here. The nature, scope, and origins of the phenomenon may be controversial, but its existence is not. And the existence of the Unintentional Fallacy is all that we need to answer the objections to the Simple View.

6. The Objections Reconsidered

Let us return to Harman’s sniper and focus on the following statements:

(10) The sniper alerted the enemy intentionally.
(11) The sniper did not alert the enemy intentionally.
(12) The sniper alerted the enemy unintentionally.

The negation of (10) is logically equivalent to (11). However, the denial of (10) will typically communicate (12), owing to the Unintentional Fallacy. And (12) is obviously false in the case described, since the relevant effect was clearly anticipated (and thus not accidental). Speakers wish to avoid communicating falsehoods, so there is pressure to avoid denial. That generates corresponding pressure towards assertion. Noting that speakers routinely mistake assertability for truth, the defender of the Simple View will suggest that the Unintentional Fallacy is to blame for our incorrect intuition on (10).

(The defender of the Simple View need not suggest that this line of reasoning consciously occurs to those who judge (10) false. Nor should they. When a trustworthy source testifies against (1), one takes the truth of (3) as given. This move is immediate, unconscious, and often irresistible—even for those who recognize the possibility of (2). What goes for (1) and (3) goes for (10) and (12).

Turning to Knobe’s Harm Vignette, we have:

(13) The chairman harmed the environment intentionally.
(14) The chairman did not harm the environment intentionally.
(15) The chairman harmed the environment unintentionally.

Once again, the defender of the Simple View will insist that the negation of (13) is logically equivalent (14). However, the denial of (13) will typically

---

20 The interested reader is directed to Horn (2001, chapter 2).
communicate (15), which is to say that the harm was accidental. This implication is false, but it is also dangerous since we do not normally hold people responsible for accidents. The chairman is clearly deserving of blame, so there is significant pressure to avoid the suggestion of (15). Hence, there is significant pressure to avoid the denial of (13). It is no surprise that 82% of Knobe’s respondents opt for affirmation instead.

Moving to the Help Vignette, we have:

(16) The chairman helped the environment intentionally.
(17) The chairman did not help the environment intentionally.
(18) The chairman helped the environment unintentionally.

As in the previous cases, the denial of (16) will communicate (18), which is false. We should therefore expect some corresponding pressure to assert (16). Sure enough, almost a quarter of Knobe’s respondents judge (16) to be true. But what explains the asymmetry between the Harm and Help Vignettes? In both cases denial would communicate something false, but in the latter case the falsehood would be largely innocuous. After all, the chairman does not deserve praise for helping the environment since he did not bring about the good result for the right reason. In contrast, the chairman does deserve blame in the first case, so it is much more important to avoid the suggestion of an accident. Moral considerations thus intensify the semantic presupposition or pragmatic forces that are already in place.

Moving on to the problem of luck, we have:

(19) Ned made the basket intentionally.
(20) Ned did not make the basket intentionally.
(21) Ned made the basket unintentionally.

This case raises two questions. First: If the Simple View is correct, why do some people affirm (19)? After all, Ned did not believe—and therefore did not intend—that he would make the shot. As in the case of side-effects, the defender of the Simple View will suggest that the intuitions threatening his position are the byproduct of understandable but fallacious reasoning. To deny (19) is to communicate (21), which is clearly false since Ned was trying to make the shot. Hence, there is some pragmatic pressure towards assertion.

The second question is more interesting: Why is there so much ambivalence on this issue? Some speakers are inclined to affirm (19), some are inclined to deny it, and some have no clear intuition one way or the other. Why? The answer, I think, has to do with the details of the case and the background knowledge shared by speaker and audience. Basketball
players put the ball in the hoop. That is their job. When Ned has the basket in plain view and throws the ball in the direction of the hoop, it is obvious what he is trying to do. When the ball goes in, there is no temptation to view this as an accident. This fact mitigates the semantic presupposition or pragmatic forces behind the Unintentional Fallacy. When one denies (19), there is less danger of communicating (21). Since the risk is low, the pressure is off. The result is that many speakers are ambivalent.

Things are very different in Butler’s case:

(22) Brown shot Smith intentionally.
(23) Brown did not shoot Smith intentionally.
(24) Brown shot Smith unintentionally.

Here, the denial of (22) communicates (24), which is not only false, but dangerous. Brown is clearly responsible for the shooting, since that was what he was trying to do. Since we do not hold people responsible for accidents, we must avoid the suggestion of (24). Hence, there is little ambivalence about (22).

Recalling the problem of consistency, we have:

(25) The player hit Target 1 intentionally.
(26) The player did not hit Target 1 intentionally.
(27) The player hit Target 1 unintentionally.

Bratman suggests that (25) is true, despite the fact that he had no prior intention to hit the target in question. He concludes that the Simple View is false. My response should come as no surprise. The denial of (25) would communicate (27), which is clearly mistaken. Hence, there is pressure to assert (25) instead.

This response is nothing new, so let me add one further observation. In his discussion, Bratman regularly moves back and forth between (25) and (28):

(28) The player intentionally hit Target 1.

Bratman treats these statements as interchangeable, but I detect a subtle difference—to my ear, (28) sounds worse than (25). Why? Here is one possible explanation. Word order can affect linguistic focus, since emphasis is often placed on the final word or clause in a sentence. Thus, it is natural to “hear” (25) and (28) as:
Now: Why should the shift in focus matter? Hypothesis: When one is asked to rule on (25*) and (28*), the shift in emphasis semantically enforces (or pragmatically implies) a different range of answers from which one is to select. These options are represented in different contrast clauses implicitly attached to (25*) and (28*):

(25**) The player hit Target 1 intentionally rather than not.
(28**) The player intentionally hit Target 1 rather than some other target.

Now we have a complete explanation for why (28) sounds worse than (25). It is natural to interpret (28) as (28**). That statement is clearly false, since the player has no special interest in Target 1. Hence, we are more inclined to deny (28). In contrast, it is natural to interpret (25) as (25**), where the focus is on intentionality itself. Here, the looming threat of the Unintentional Fallacy pushes us towards assertion.

7. Intentional Action and Moral Responsibility

I have suggested that the Unintentionally Fallacy underlies all of the objections to the Simple View of intentional action. If I am correct, all of those objections are simply mistaken.

In closing, I would like to draw a more general lesson about intentional action and moral responsibility.

It is sometimes said that one can only be held responsible for actions that one performs intentionally. This idea plays an important role in many legal systems, most notably in the mens rea requirement for legal culpability. The concept of mens rea traces back to the writings of Henry Bracton in the thirteenth century, who said that “it is will and purpose which mark maleficia… and crime is not committed unless the intention to injure exists.”21 Courts today make exceptions in cases of recklessness and negligence, but Bracton’s core idea remains intact: If you don’t do it intentionally, you aren’t responsible. This is the standard view.

The standard view is endorsed, in one form or another, by many philosophers. Sidgwick, for example, says that:

---

21 Bracton (1250, p.290).
For the purposes of exact moral or jural discussion, it is best to include under the term ‘intention’ all the consequences of an act that are foreseen as certain or probable, since... we cannot evade responsibility for any foreseen consequence of our acts by the plea that we felt no desire for them.22

Sidgwick’s conclusion is that desire is not necessary for intentional action. His argument is that, if it was, it would be too easy to evade responsibility. Implicit in this argument is the assumption that you are not responsible if you do not act intentionally.23

I think Sidgwick and Bracton are both mistaken. The standard view is false, even if we qualify it to take account of recklessness and negligence. In fact, the view is undermined by Sidgwick’s own example, as well as those discussed by Butler and Knobe.

Sidwick says that you cannot avoid responsibility for a foreseen effect by saying that you felt no desire for it. I agree. He concludes that we must redefine “intention” so that it is synonymous with “foresight”. I disagree. The correct response is to reject the Standard View and sever the presumed connection between intention and responsibility.

In fact, the standard view is one more mistake caused by the Unintentional Fallacy. The correct view is this: If one does something unintentionally, one isn’t responsible. The correct view is correct, at least as an approximation. The problem is that the correct view is mistakenly equated with the standard view because of our tendency to treat contraries

---

23 The same assumption is made by Anscombe in discussing the “Cartesian” view that “an intention was an interior act of the mind which could be produced at will.” (1981, p.59) She complains that

on this theory of what intention is, a marvellous way offered itself of making any action lawful. You only had to ‘direct your intention’ in a suitable way. In practice, this means making a little speech to yourself: “What I mean to be doing is...” ...This same doctrine is used to prevent any doubts about the obliteration bombing of a city. The devout Catholic bomber secures by a ‘direction of intention’ that any shedding of innocent blood that occurs is ‘accidental’. (ibid., p.59)

Anscombe’s idea is clear. If the death of the civilians was not the intention of the bomber, then those deaths were unintended and hence ‘accidental’. Since those deaths were clearly not accidents, Anscombe concludes that they must have been intended.
like contradictories. We mistakenly assume that, if something is not intentional, it is unintentional. And if it is unintentional, you cannot be held responsible. So, you cannot be held responsible if it is not intentional. That line of reasoning is attractive but unsound. If I am correct, there is an important gap between unintentional and non-intentional. Those who fail to mind the gap will miss out on the true relationship between intentional action and moral responsibility.

References