INTENDED AND FORESEEN UNAVOIDABLE CONSEQUENCES

DEVLIN RUSSELL

York University
4700 Keele Street, Toronto, Ontario M3J
Canada
devlinrussell@fastmail.com

Article info
CDD: 153.4
Received: 03.09.2018; Revised: 09.10.2018; Accepted: 15.10.2018

Keywords:
Intention
Intentional action
Reasons for action
Cognitivism about intention
Deviant causation

Abstract: What is the difference between an intended consequence and a foreseen unavoidable consequence? The answer, I argue, turns on the exercise of knowhow knowledge in the process that led to the consequence. I argue for this using a theory according to which acting intentionally is acting as a reason. I show how this gives us a more promising explanation of the difference than the dominant explanations, according to which acting intentionally is acting for a reason.

A bomb has fallen and destroyed a munitions factory, killing twenty innocent civilians. Here is what we know. The bomber was a perfectly rational agent. She deliberated about whether to drop the bomb and knew that dropping the bomb would both completely and irreparably disable the last
remaining munitions factory of an evil enemy and kill twenty innocent civilians who were working in the factory unaware of the factory’s true function. Oddly but with good reason in her mind, she valued the destruction of the factory on exact par with the death of the innocent civilians. She weighed, without error, the cost and benefits of dropping the bomb, determining that dropping the bomb was the best option, and as a result decided to drop it. Everything from deliberation to destruction proceeded in a perfectly normal way.

Here is what we don’t know: whether or not the death of the innocent civilians was an intended consequence of the bombing. We know that the bomber wanted to end the war, and we know that she knew that destroying the factory would be an effective way to end the war because it would be an effective way to significantly disable the enemy and force them to give up. It turns out, however, that killing the innocent civilians would be an equally effective way to end the war because it would be an equally effective way to terrorize the enemy and force them to give up. And the bomber knew this. Also, the bomber valued disabling the enemy on exact par with terrorizing the enemy. So it turns out that the bomber knew that ending the war was overdetermined (two courses of action were equally effective at ending the war), and those two courses of action were, for her, tied for best option. However, despite knowing all this, we do not know whether she intended to end the war by destroying the factory (thereby disabling the enemy) or to end the war by killing the innocent civilians (thereby terrorizing the enemy). She knew that the destruction of the factory and the death of innocent civilians were both unavoidable consequences of dropping the bomb, but which consequence was intended?

As philosophers, we are not looking for more evidence about the case but inquiring into the nature of the phenomenon itself. What is the difference between an intended consequence and a foreseen unavoidable consequence? What is the difference between two possible cases that are exactly
alike in every respect except in the one case the consequence (e.g., the death of innocent civilians) is intended while in the other the consequence is not intended but foreseen as unavoidable.

We want to understand this difference not just because it has important moral and legal implications but also because it is utterly commonplace. Everything that human agents do in this world have unavoidable, well-known, perfectly benign consequences. All of them cannot be intended. Understanding why can help us understand human agency.

In this paper, I will argue that a theory according to which acting intentionally is acting as a reason can help us give a more promising explanation of the difference between our two cases than a theory according to which acting intentionally is acting for a reason.

1. THE PUZZLE

According to a popular non-cognitivist theory of intention, an intention to bring about a consequence is a sui generis, conative mental state without any cognitive commitment to whether one will bring about that consequence (Mele 1992; Bratman 1999). For example, when an agent has an intention to bring about a loaf of bread, she is motivated in certain ways to bring about a loaf of bread, but her intention does not involve some commitment to the truth that she will bring about a loaf of bread. According to a popular, rival theory, an intention to bring about a consequence does involve a cognitive commitment. In particular, it involves the belief that one will bring about that consequence by this very belief. Thus, when an agent has an intention to bring about a loaf of bread, her intention is a belief that she will bring about a loaf of bread by this very belief (Harman 1976; Velleman 1989, Setiya 2007).
Both the cognitivist (of this stripe) and the non-cognitivist (of this stripe) agree that an intention is, characteristically, a conclusion of deliberation. That is to say, an intention is, characteristically, formed by weighing the reason for and against bringing a consequence about and resolving to bring that consequence about.

Against cognitivism, some non-cognitivists argue that this cognitivism cannot explain the difference between our two cases. Their reasoning goes like this. In both cases, the bomber deliberated about whether to bring about the bombing and resolved to bring about the bombing. On the cognitivist view, therefore, the bomber intended to bomb and thus believed that she will bring about the bombing by this very belief. However, the conclusion of her deliberation does not seem to be merely that she will bring about the bombing by this very belief but also that she will bring about the destruction of the factory by this very belief and that she will bring about the death of innocent civilians by this very belief. After all, she knew that her belief would bring about not just the bombing but also the destruction of the factory and the death of innocent civilians, and she determined that was the best option. Therefore, on the cognitivist view, it looks like, in both cases, the bomber did not just intend to bring about the bombing but also to bring about the destruction of the factory and the death of innocent civilians. (Bratman 2009)

The cognitivist view cannot, therefore, explain the difference between the two cases because the cognitivist view implies that both the destruction of the factory and the death of innocent civilians were intended consequences. But the difference between the cases just is that, in the one, the destruction of the factory is intended while the death of innocent civilians is not and, in the other, the destruction of the factory is not intended while the death of innocent civilians is.
In order to explain the difference, the cognitivist needs to explain what it is that makes the relevant consequence not intended but merely a foreseen, unavoidable consequence.

One attempt to meet this challenge goes like this: the difference between the two cases is in some conception that the agent has of the instrumental connection between the various consequences. In one case, the bomber conceives of the destruction of the factory as a means to ending the war but does not conceive of the civilian deaths as a means to ending the war. In the other case, it is the other way around. So although the agent believes that the relevant consequence will, unavoidably, be brought about by her belief, she does not conceive of that consequence as a means. This explains why the consequence is not intended. (Marusic & Schwenkler, forthcoming)

Unfortunately, this attempt is not explanatory. How are we to understand the agent’s conception of the instrumental connection? If her conception is how she deliberates about whether to bomb, then her conception looks the same in both cases, and thus this response implies that there is no difference. If her conception is a belief about the effects of the various consequences, then again her conception looks the same in both cases. In both cases, she believes that the bombing will (unavoidably) bring about the destruction of the factory and the death of the civilians, which in turn will (unavoidably) bring about the disablement and the terror of the enemy.

If her conception is the belief that the bombing is a means to the destruction of the factory, in some special sense of ‘means’ that is not equivalent to ‘something that will bring about’, then this response is in danger of circularity. For the natural, stronger reading of ‘means’ is ‘something that will bring about an effect intentionally’. Thus, the response amounts to saying that the difference (in our kind of case) between an intended consequence and a not intended one is that the former is brought about intentionally while the other
is not, which amounts to giving no explanation at all. In order to explain the difference, the cognitivist needs to explain the difference between a means and something that brings about an effect non-intentionally. (Paul 2011)

At this point, the non-cognitivists argue that their theory can help us give a more promising explanation than cognitivism of the difference between our two cases. According to these non-cognitivists, an intention to bring about a consequence does not consist in certain cognitive commitments but in certain conative commitments. In particular, it consists in commitments to follow through with bringing about the consequence (e.g., to not reopen deliberation, to reason about how to bring the consequence about, and to execute the results of that reasoning until complete). The difference between our two cases lies in the bomber’s commitment to follow through with bringing about the relevant consequences. In the one case, the bomber was committed to following through with bringing about the destruction of the factory, but she was not so committed regarding the death of the civilians. In the other case, it was the other way around. The relevant consequence was intended because it was brought about by such a commitment, while the other consequence was not intended because the bomber was not so committed, even though she knew it was an unavoidable consequence of her commitment to bomb. (Bratman 2009)

Given this explanation, it seems that these non-cognitivists are open to an objection structurally parallel to theirs against the cognitivists. In both cases, the bomber deliberated about whether to bring about the bombing and resolved to bring about the bombing. On this non-cognitivist view, therefore, the bomber intended to bomb and thus was committed to following through with bringing about the bombing. However, the conclusion of her deliberation does not seem to be merely a commitment to following through with bringing about the bombing but also a commitment to following through with bringing about the destruction of the
factory and a commitment to following through with bringing about the death of innocent civilians. After all, she was committed to following through with bringing about the bombing, knowing full well that the bombing would (unavoidably) bring about both the destruction of the factory and the death of innocent civilians. So she seems committed to following through with bringing about the death of innocent civilians as much as she’s committed to following through with bringing about the destruction of the factory. Therefore, even on this non-cognitivist view, it looks like, in both cases, the bomber did not just intend to bring about the bombing but also to bring about the destruction of the factory and the death of innocent civilians.

But, these non-cognitivists will say, the bomber was not committed to following through with bringing about the death of innocent civilians in both cases. This is because, in the one case, she would have done nothing differently, were she to have discovered that her bombing would not successfully bring about the deaths, while in the other, she would have reasoned anew about how to bring about the deaths. This is part of what the bomber’s commitment consists in.

So, the relevant consequence was intended because it was brought about by such a commitment, while the other consequence was not intended because the bomber was not so committed, even though she knew it was an unavoidable consequence of her commitment to bomb. The problem with this explanation is that, although it identifies a difference in the bomber, this doesn’t seem to be the difference that makes a consequence intended or not. Suppose all else was equal, and the bomber was committed to bringing about the destruction of the factory but not to bringing about the death of the civilians. And suppose further that forming this commitment brings about a state of nervousness in the bomber, which in turns brings about the tensing of her muscles, which in turn brings about a triggering of the bomb. Therefore, in the end, her commitment to bringing about the
destruction of the factory brought about the destruction of the factory, but the destruction was not intended. The bomber’s commitment brought about the consequence, but the consequence was not brought about with that intention. It was not brought about with any intention at all. Thus, even in our original cases, something more than the bomber’s commitment must explain why she intended the relevant consequence. (Wilson 1989; Sehon 1997)

One response is to say that the relevant consequence must be brought about by the sustained guidance of the commitment. In our supposed counterexample, guidance was not sustained. At some point, the nervousness interrupted the guidance of the commitment (Frankfurt 1978). However, this response is too strong. In our original cases, the intended consequences were not brought about by the sustained guidance of the commitment. When the destruction of the factory was intended, it is possible that the commitment to follow through with bringing about the destruction of the factory guided the process up to the moment of triggering the bomb, but beyond that it is not possible that the commitment had any causal influence over the process. So, the destruction of the factory was intended, yet it was not brought about by the sustained guidance of the corresponding commitment.

If the non-cognitivist responds that there is some special, appropriate way that a consequence is brought about when it is intended, then this response is in danger of circularity. For to say ‘some special, appropriate way’ is simply to say ‘the way that makes the consequence intended’. Thus, the response amounts to saying that the difference (in our kind of case) between an intended consequence and a not intended one is that the former is brought about by a certain commitment in a way that makes the consequence intended while the other is not, which amounts to giving no explanation at all. In order to explain the difference, the non-cognitivists need to explain the difference between bringing about
a consequence in a way that makes it intended and merely bringing it about.

Therefore, the non-cognitivists’ explanation of the difference seems no more promising than the cognitivists’. Both leave us wanting.

None of the arguments above are decisive against cognitivism or conative-state-based non-cognitivism in general. There are a number of strategies that proponents could employ to save their views, though many of these will make the explanation of the difference between our cases far less natural. In what follows, I will argue against trying to save these views by arguing for the promise of an explanation that rejects them both.

2. Diagnosis of the Puzzle

Although both the explanations we’ve consider so far are initially plausible and laudable, I suggest that they suffer from a common error. Namely, they both assume that an intended consequence is one brought about by an agent’s intention. This naturally leads to the thought that an intention is a state of the agent—namely, the state that brings about a consequence when that consequence is intended. Thus, they try to explain the difference between an intended consequence and one that is not in terms of the state of the agent. The cognitivists claim that the agent has a special sort of belief about the consequence when it is intended. The non-cognitivists claim that the agent has a special sort of conative commitment to the consequence.

I suggest that I can give a more promising explanation than these cognitivists and non-cognitivists by denying that an intended consequence is one brought about by an agent’s intention. What is the alternative?

I propose that an intended consequence is one brought about intentionally, and a consequence was brought about...
intentionally just in case this bringing about was the agent’s reason for bringing some other consequence about. Therefore, a reason for bringing about a consequence is not a state of the agent that brings about the consequence. Rather, it is a wider process of bringing something about, part of which is a narrower process of bringing about the consequence. Therefore, explaining the difference between our two cases will consist in explaining the difference in the processes of bringing about the relevant consequences.

The explanation can be simply put, but will need elaboration. The difference between the two cases lies not in the states of the bomber but in the parts of the process that brought about the consequences. In both cases, the bomber had all the same knowledge and knowhow, but the relevant consequence was intended only when part of the process that brought it about was an exercise of knowledge how to bring it about by such an exercise.

3. ELABORATION OF THE SOLUTION

Let us start by observing that wants, intentions, attempts, and doings are all, characteristically, kinds of reasons for action. Here I do not mean ‘reason for action’ in the sense of being a reason that counts in favour or against performing an action. Rather I mean ‘reason for action’ in the sense of that for the sake of which the agent acted. In this sense, I can raise my arm because I want to catch a ball or because I intend to catch a ball or because I am trying to catch a ball or because I am catching a ball. Put another way, I can bring about an arm-movement because I want to bring about a catching or because I intend to bring that about or because I am trying to bring that about or because I am bringing that about. It is no accident that all these things are reasons in this sense.
According to my developmental theory of action (Russell 2018), want, intention, attempt, and doing are, respectively, analogous to bud, blossom, set, and (mature) fruit in ways that help explain the nature of reasons for action.

Buds, blossoms, sets, and fruits are all, characteristically, kinds of explanans. They explain their parts in terms of a whole. Some bit of matter might be arranged in a certain way because it is (in part) a bud or is (in part) a blossom or is (in part) a set or (in part) a fruit. According to the developmental theory, it is the same with wants, intentions, attempts, and doings. They are explanans of explanations in terms of a whole. In the above examples, the bringing about of the arm-movement is part of the want or intention or attempt or doing. This implies that wants, intentions, attempts, and doings are not states of an agent but processes of bringing a consequence about. In particular, a want to bring about a consequence, C, is (among other things) a bringing about of C. And an intention to bring about C is (among other things) a bringing about of C. For example, a want or intention to bring about a caught ball is (among other things) a bringing about of a caught ball. The arm-movement in the above example is a part of a bringing about of a caught ball. The characteristic explanatory relation between wants, intentions, attempts, and doings, on the one hand, and the bringing-about they explain, on the other, is the relation of whole to part.

Thus, a want to bring about C, an intention to bring about C, an attempt to bring about C, and a bringing about of C have one important feature in common: each is (among other things) a bringing about of C.¹ They are different because they are at different characteristic stages of develop-

¹ Note that the existence of a bringing about of C does not entail that C was or will be brought about. See Falvey 2000 and Thompson 2008 (especially Part II).
ment. A bud is a fruit at a characteristic stage of development. A blossom is a fruit at a characteristic stage of development. In the same way, a want is a bringing-about at a characteristic stage of development. And an intention is a bringing-about at a characteristic stage of development. A doing, like a mature fruit, is a bringing-about at the mature stage of action-development.

What distinguishes a bud, a blossom, a set, and a (mature) fruit from each other is the developmental role it plays. A bud’s role is to grow certain parts of the blossom. A blossom’s role is to achieve fertilization. And so on. Similarly, according to the developmental theory, a want, an intention, an attempt, and a doing are distinguished from each other by the developmental role each plays. The specifics are not essential for the purposes of this paper, but some suggestions will give us a better picture of the view. Roughly, the role of a want to bring about C is to determine whether to bring about C. The role of an intention to bring about C is to determine how to bring about C and to plan and prepare for an attempt to bring about C. The role of an attempt to bring about C is to bring about the consequences determined at the intention-stage. And the role of a mature bringing about of C is to complete the bringing about of C.

Suppose that I am examining an apple because I want to eat it. I am in the process of eating the apple, or (i.e.) bringing about an eaten apple, and my examining is a part of that process. But my apple-eating is not a mature apple-eating (when we would say I am eating it) nor is it an intention to eat an apple. My examining functions to determine whether to eat the apple, and thus my apple-eating is in the developmental stage of being a want. Suppose that I am flipping through my day planner because I intend to meet someone for lunch. I am in the process of meeting someone for lunch, or (i.e.) bringing about that meeting, and my flipping is a part of that process. But my meeting is not a mature meeting (when we would say I am meeting the person for lunch) nor is it a mere
want to meet the person for lunch. My flipping functions to plan on meeting the person for lunch, and thus my meeting is in the developmental stage of being an intention. And so on for attempts and doings.

But what makes one bringing-about a part of another and play a specific role? After all, I can want to make a fruit salad and be examining an apple, but the examining is not part of the salad-making. Perhaps the apple is a snack for now, while the fruit salad is for guests later. Further, I can be examining an apple to determine where to bite it, thus intending to eat it rather than merely wanting to. What makes it the case that I am examining the apple in order to eat it and that this examining is functioning to determine whether to eat it?

It is the exercise of instrumental knowhow. Instrumental knowhow is not knowhow knowledge of, simply, how to bring something about. It is, more specifically, knowhow knowledge of how to bring something about by bringing something else about. When I exercise instrumental knowhow of how to determine whether an apple is edible by examining an apple, I not only examine an apple but also determine whether an apple is edible, and the former is a part of the latter. In general, instrumental knowledge is knowhow knowledge of how to bringing about a consequence, E, by bringing about some other consequence, M. To exercise instrumental knowhow is to bring about M as part of a bringing about of E. Therefore, if an agent brought about M and brought about E, and the bringing about of M was an exercise of instrumental knowhow, then her bringing about of M was a part of her bringing about of E.

If all of the parts of the agent’s bringing about of E taken together were sufficient to bring about E, then her bringing about of E was intentional (i.e., she brought about E intentionally). The converse is also true. This is the mereological theory of intentional action. And if her bringing about of E was intentional (in this way), then it is her reason for bringing
about M. The converse is also true. This is the acting-as-a-reason theory: an intentional action is an action performed as a reason, not for a reason. And finally, if her bringing about of E was intentional, then E was an intended consequence (and conversely). This final claim is common ground.

We can now explain the difference between our two cases. In the one case, the bomber brought about a triggering and brought about the destruction of the factory, and the bringing about of the triggering was an exercise of knowhow knowledge of how to bring about the destruction of the factory by bringing about the triggering. And the triggering was sufficient to bring about the destruction. Therefore, her triggering was a part of her bringing about of the destruction of the factory, and it was sufficient. Thus, her bringing about of the destruction of the factory was intentional, and, in turn, the destruction of the factory was intended. Her triggering was not a part of her bringing about of the deaths because the triggering was not an exercise of knowhow knowledge of how to bringing about the deaths by bringing about the triggering. And thus the deaths were not an intended consequence, even though she knew that she will bring about the deaths and knew how to bring them about by triggering. In the other case, it is the other way around.

To put it briefly, the intended consequence is intended because the agent’s instrumental knowhow was exercised as a sufficient part of the process that brought it about.

And thus, the difference between an intended consequence and a non-intended one lies not in what brought about the process that led to it but in the modal properties of this process. Exercises of knowhow manifest non-accidental counterfactual success, parallel to non-accidental true belief (Hawley 2003). It is no accident in this sense that (e.g.) the factory was destroyed by the triggering. By contrast, the

---

2 See Thompson 2008, especially Part II, and my Russell 2018 for further defence.
deaths of the innocents were an unavoidable, accidental success (because the relevant knowhow was not exercised and thus its modal properties were different).

4. OBJECTIONS AND REPLIES

Suppose that I brought about the word ‘action’ by bringing about the depression of the ‘a’ key, and the bringing about of the depression was an exercise of my knowhow knowledge of how to bring about the word ‘action’ by bringing about the depression of the ‘a’ key. But here’s what happened. I pressed the ‘a’ key, my computer glitched, and it randomly outputted the word ‘action’. It seems that, according to my explanation, the word ‘action’ was an intended consequence because my instrumental knowhow was exercised as a sufficient part of the process that brought it about. However, the word ‘action’ was not intended here because the bringing about of it was not intentional.

My reply is to refine the mereological theory of intentional action. It is a simplification to say that I knew how to bring about the word ‘action’ by bringing about the depression of the ‘a’ key. What I know, more precisely, if I know this, is how to bring about the word ‘action’ by bringing about the depression of the ‘a’ key, followed by the depression of the ‘c’ key, followed by the depression of the ‘t’ key, and so on. If an agent knows how to bring about E by bringing about M, she knows how to bring about E by bringing about M as part of what’s sufficient for bringing about E. More precisely, then, a consequence, E, was brought about intentionally just is case all of the parts of the agent’s bringing about of E taken together were sufficient to bring about E and corresponded to the complete and precise specification of the instrumental knowledge exercised.

In the above example, although the bringing about of the depression of the ‘a’ key was an exercise of my knowledge.
how to bring about the word ‘action’ by bringing about the depression of the ‘a’ key and thus was a sufficient part of bringing about the word ‘action’, the parts of the bringing about of the word ‘action’ did not correspond to the complete and precise specification of the knowhow exercised in bringing about the depression of the ‘a’ key.

Now suppose that I brought about a bullseye by bringing about an arrow-shot, and the bringing about of the arrow-shot was an exercise of my knowhow knowledge of how to bring about a bullseye by bringing about that arrow-shot. But here’s what happened. I shot the arrow, a westward wind blew the arrow off course, then an eastward wind blew it back on course. It seems that, according to my explanation, the bullseye was an intended consequence because all of the parts of my bringing about of the bullseye taken together were sufficient to bring about the bullseye and corresponded to the complete and precise specification of the instrumental knowledge exercised. Shooting the arrow in just the way I did was sufficient to hit the bullseye. This example is not like the typing example, where I did not type what I knew to type. Rather, in this example, I did everything I knew to do. However, my bullseye was not intended because it was not brought about intentionally.

This objection reveals some imprecision in my explanation. My response is to be more precise. The instrumental knowhow that is exercised in intentional action must be specific to the situation, not cross-situational. In my explanation, I claimed that instrumental knowledge is knowhow knowledge of how to bringing about a consequence, E, by bringing about some other consequence, M. Cross-situational instrumental knowledge is knowhow knowledge of how to bringing about E by bringing about M, across many, various situations. If an agent has cross-situational instrumental knowledge of how to bring about a loaf of bread by bringing about some dough, she could exercise this knowledge at many different times, in many different places, and (more
generally) under many different conditions. But if an agent has situationally specific instrumental knowledge of how to bring about a loaf of bread by bringing about some dough, she can exercise this knowledge only under very limited conditions close to her present situation. Situationally specific instrumental knowledge is knowhow knowledge of how to bringing about E by bringing about M, in situations very much like the one under evaluation. It is this sort of instrumental knowledge that must be exercised in order to bring about M as a part of a bringing about of E.

In the arrow-shot example, I did not have situationally specific instrumental knowledge of how to bring about a bullseye by bringing about that arrow-shot. I didn’t know how to do this under the present wind conditions. If I had, I would have hit the bullseye intentionally, utilizing the wind currents.

CONCLUSION

So here is what we didn’t know: whether or not the process that brought about the death of the innocent civilians consisted of parts that, taken together, were sufficient to bring about the deaths and, individually, were exercises of situationally specific knowhow of how to bring about the deaths by bringing about those parts. If, all else being relatively normal, the bomber’s triggering was not an exercise of knowledge of how to bring about the deaths by so triggering (in situations very much like this one), then the deaths were not intended.
REFERENCES


