

Stout's teleological theory of action

Jeff Speaks

November 26, 2004

1	The possibility of externalist explanations of action	2
1.1	The distinction between externalist and internalist explanations of action	2
1.2	The internalist shift	2
1.2.1	Against facts justifying actions	2
1.2.2	The argument from false beliefs	3
1.2.3	The argument from the impotence of unrepresented facts	4
1.2.4	The argument from our awareness of reasons for action	5
1.3	Stout's argument for externalist explanations	5
2	Causal and teleological explanation	6
2.1	Causal explanation	6
2.2	Teleological explanation	7
2.2.1	Why teleological explanation should count as a kind of explanation	7
2.2.2	Teleological explanation and intentional action	8
3	Results	9
3.1	Intentional action	9
3.2	Belief	11
3.3	Intention	11

Stout presents the framework of his teleological theory of action in Chapter 1:

The teleological theory of action A 's ϕ ing is an intentional action \equiv_{df} A 's ϕ ing is explainable in terms of a practical justification.

So construed, Wilson's view is also an instance of the teleological theory; he claimed that the intentional actions were the ones done in order to do something. As we'll see, Davidson's view is also a version of this theory, according to Stout. A fair conclusion is that, by itself, the teleological theory of action doesn't say that much; the interest of a version of the theory will be in how it explains the key notions in terms of which the teleological theory is defined. Stout's version of the teleological theory will end up being quite different from Wilson's and Davidson's; most of the book is devoted to spelling out exactly the version of this view that he wants to defend.

1 The possibility of externalist explanations of action

1.1 *The distinction between externalist and internalist explanations of action*

The first step in doing this (pp. 11 ff.) is to distinguish between two different kinds of teleological explanations of action, which Stout calls internalist and externalist explanations.

Roughly speaking, an internalist explanation of action is an explanation in terms of an agent's mental states. Stout's example is

I missed assembly because I believed that it would exacerbate my asthma, and
I did not want that.

If we analyze 'explainable in terms of a practical justification' as something like 'caused by a belief-desire pair connected to the action and some goal in the right way', then we get Davidson's theory as a version of the teleological theory. More generally, psychologistic versions of the teleological theory are ones which treat being 'explainable in terms of a practical justification' as a matter of having some internalist explanation.

Roughly speaking, an externalist explanation of action is an explanation *not* in terms of an agent's mental states. Stout's example is

I missed assembly because it would have exacerbated my asthma.

If, like Stout, you want to defend a non-psychologistic version of the teleological theory, you will want the kind of explainability employed by the theory to be a kind of externalist explanation.

1.2 *The internalist shift*

Stout says, correctly, that there is some tendency to think that any externalist explanation of an action is underwritten by some more fundamental internalist explanation of action. E.g., if one is presented with the claim that Bob missed the assembly because it would have exacerbated his asthma, one can ask: how could the fact that it would have exacerbated his asthma cause him to miss the assembly? And it seems that you can answer this question by giving an internalist explanation like the one cited above.

Stout's aim in the second half of Ch. 1 is to undercut some of what he takes to be the motivations for this kind of tendency toward an internalist shift.

1.2.1 *Against facts justifying actions*

One such motivation is the idea that it is impossible for facts about the external world to justify an action. Stout suggests that one source of this idea is Hume's claim that any

justification of an action has to involve a desire. He then suggests two ways of backing up this Humean claim.

1. *External facts are motivationally inert, and hence cannot justify an action.* Stout replies that external facts might be in some sense motivationally inert; all that he's claiming is that facts can give reasons for action, as in the case where the fact that you stepped on my toes, I am thereby have a reason to complain. He claims that this is analogous with belief, where the fact p by itself can justify the belief p . (It's far from obvious that this works even in the case of belief. We are very inclined to say that I must, e.g., perceive p for it to figure in a justification of the belief p .)

2. *The demand for total justification.* When you ask someone, 'Why should x count as a reason for ϕ ing?', it seems that one bottoms out in the response, 'I wanted to $_$.' But this seems to indicate that some want or other is at the foundation of all of our reasons for action. Stout replies that explanations like this never bottom out. We can ask, as he says, what justifies one's choice for being a desire-fulfiller rather than a desire-denier. So we should not say that the only real explanations are the ones not susceptible to the 'why' question, since there are no such explanations.

1.2.2 *The argument from false beliefs*

The above is a puzzle about how facts could justify actions. But there are two further puzzles about how external facts could *cause* actions.

The first of these is what Stout calls the 'argument from false beliefs.' The argument is simple enough:

1. When an agent acts on the basis of false beliefs, their actions are caused by their beliefs (rather than by the external facts corresponding to those beliefs since, the beliefs being false, there are none).
 2. This kind of causation is operative in the case where the agents beliefs are true as well.
 3. If a phenomenon can be explained in every case without reference to x , then x is not an essential part of the immediate explanation of the phenomenon.
-
- C. External facts do not figure in the most immediate explanations of action. (I.e., these sorts of explanations are always underwritten by more fundamental explanations in terms of the mental states of the relevant agent.)

Stout offers two kinds of responses to this argument: (i) he argues that accepting this argument would lead to undesirable consequences in other areas, and (ii) he suggests that we can block the argument by denying premise 3.

He offers three versions of response (i):

1. If we accept the argument, then we must reject the view that actions can be (im-

mediately) explained by factive mental states as well as the view that they can be explained by external facts. *This is unlikely to upset anyone who accepts the argument.*

2. If we accept the argument, then we must accept, by parity of reasoning, the conclusion of the Argument from Illusion. This would mean believing that only facts about how things appear to us, rather than facts about how things are, can be a part of the most fundamental justification for perceptual beliefs. *Usually the Argument from Illusion is presented as an argument for the sense datum theory of perception, which does seem to be a view worth blocking. But the version of the Argument from Illusion that Stout presents does not seem particularly 'dangerous.' Isn't it true that our most immediate reasons for our perceptual beliefs are things about how the world perceptually seems to us? What is supposed to be wrong with that view?*
3. If we accept the argument, then we have to think of intentional actions as internal acts of effort, rather than external events which involve interaction with other things. After all, no action which extends into the external world can be explained wholly in terms of an internalist explanation – facts about the external world must be involved as well. *The argument does not appear to be quite this easy. Some explanations are good, useful, true, etc. without being complete. Surely an internalist explanation of the sort being discussed would not be the full and complete explanation of any intentional action, since such an action rests on external factors as well. But what is wrong with that? More to the point, are externalist explanations of action any different? Do they express every fact on which the action in question depends? Perhaps there is a better way of running this argument than I've stated; but I don't think that I see what it is.*

Stout's positive suggestion is that we should block the argument by denying premise 3. His idea is that the argument rests on the idea that both internalist and externalist explanations of action are on the same level; once we grant this, then it does seem unavoidable that the internalist explanations are the more fundamental ones. But we can deny this claim. This seems to me like the right kind of response to the argument; but we will have to wait on Stout's theory of explanation to see whether it is satisfactory.

1.2.3 *The argument from the impotence of unrepresented facts*

The root argument here is that, as Stout says, a fact must be mentally represented by an agent in order to be a part of the explanation of that agent's action. (Where explanation is not a matter of mere causal explanation, but rather reason-giving explanation.) Stout discusses Hume's version of this idea, which focuses on desire rather than mental representation in general.

Stout says that we can agree that there is a conceptual connection between acting for a reason and desire without thinking that desires are efficient causes of action. Rather, perhaps we can explain desire in terms of reasons for action in a noncircular way, if those reasons are thought of as being given by externalist explanations. (This is an instance of a general strategy employed by Stout: try to explain conceptual/necessary connections

between mental states and action in terms of a behaviorist account of those mental states rather than a psychologistic account of action.)

1.2.4 The argument from our awareness of reasons for action

One might argue that externalist explanations cannot give a subject's reasons for action, since we typically have a kind of immediate awareness of our reasons for action, but never have that kind of immediate awareness of external facts.

Predictably, Stout denies the claim that we never have a kind of immediate awareness of external facts; he says that he will explain immediate awareness in terms of reasons for action, rather than the reverse.

Connection to Setiya's discussion of the explanation of necessary truths; teleological behaviorism as an alternative explanation of these necessary truths.

1.3 Stout's argument for externalist explanations

In addition to trying to rebut the various arguments against the possibility of externalist explanations of action, Stout also provides one positive argument for such explanations (pp. 36 ff.). A more explicit version of it is as follows:

1. Strong internalism is false; hence either externalist explanations of action are possible or explanations of action in terms of factive mental states are possible. (Assumption)
2. Externalist explanations of action are not possible. (Assume for reductio)
3. Some actions have their most immediate explanation in terms of some factive mental state. (Since this will be a propositional attitude, we can represent an agent's being in that mental state as the agent's being in a certain relation R to some proposition p .) (1,2)
4. There must be some analysis of what it is for an agent to bear R to some proposition p .
5. If bearing R to p could be analyzed as the conjunction of a strongly internal mental state with some external fact (e.g. p itself), then the external fact would be redundant, and the explanation would be in terms of the strongly internal mental state alone.
6. Bearing R to p cannot be analyzed as the conjunction of a strongly internal mental state with some external fact. (3,5)

(6) states the preliminary conclusion of the argument. We can use it to argue for externalist explanations of action as follows: given (4) and (6), we know that the analysis of bearing R to p must take one of the following two forms: (i) it must be analyzed in terms of some external fact along with some *non*-strongly internal mental state, or (ii) it must be analyzed simply in terms of some external fact explaining some behavior of the agent.

If the analysis were of type (ii), it would follow that externalist explanations of action are possible. If the analysis were of type (i), we could then ask: how do we analyze this new, non-strongly internal mental state? And then we could run the same argument again. Either we bottom out in externalist explanations of action at some point, or we have an infinite regress.

One obvious problem with this argument is that we do not seem to have, contra what Stout seems to think, any decent argument against strong internalism on the table yet.

A second obvious problem is the assumption that any factive mental state is analyzable. Especially when one considers states like being aware that p , this does not seem especially obvious.

For these two reasons, it is hard to see this argument as very convincing.

2 Causal and teleological explanation

In Chapters 2 and 3 of the book, Stout aims to present a model of causal explanation in general which will have the result that teleological explanation is a special case of causal explanation. This lays the groundwork for the theories of practical reason, action, and mental states in Chapters 4 and 5.

2.1 Causal explanation

We will not spend much time on Stout's theory of causal explanation; we will only introduce as much of the machinery as is required by his discussion of teleological explanation. Following are some of the main theses:

- To explain why something happened is to give a causal explanation of the thing. Teleological explanation is a special case of causal explanation. There are explanations which are not causal, but explanations of action are not among them.
- A central part of an account of causal explanation is recognition of *processes* as a metaphysical category distinct from the category of events. Processes persist through time, and are not composed of temporal parts.
- We should think of the occurrence of a process as consisting in (i) a certain set of underlying conditions associated with the process being the case, and (ii) the occurrence of a certain structure of stages of the process. (Stout claims on p. 56 that the underlying conditions should be so specified that (i)'s obtaining entails that (ii) obtains, but this is dispensable.)
- There is an intuitive problem about giving an account of something in terms of whether it is explainable in a certain way. For either (following Lewis) we think of causal explanation as a matter of giving some information about the causal history of an event, in which case the notion of being explainable is too loose; or (following van Fraassen) we think of causal explanation as, like explanation generally, a pragmatic

matter dependent on what one's interest in the explanation is, in which case we cannot use explainability to define an objective phenomenon. Stout suggests that this can be solved by invoking processes. To explain why something happened is to locate that thing as part of some process; we ignore facts about why that process is happening. (78)

- There are two kinds of process-based explanations of why something is happening, which correspond to the two characterizations of what a process is. One can either locate an event within a characteristic structure of stages, or give information about the conditions which underly those stages. The former is a kind of superficial knowledge, but still should count as knowledge of why something happened. ('Why did the calculator output the number 8?' 'Because 3 and 5 were entered into it.' This locates an event within the structure of stages characteristic of a calculator doing sums.)
- Macro-level processes are typically not reducible to a series of micro-processes, just as ordinary objects are not typically reducible to the subatomic particles of which they are composed.

2.2 *Teleological explanation*

To give a teleological explanation of something is to explain it by pointing out that it is a means to some end.

2.2.1 *Why teleological explanation should count as a kind of explanation*

The puzzle about teleological explanation is to say why showing that something is a means to an end should count as an explanation of it. As Stout says, it appears to be a kind of odd backward causation: how could some end go back in time to determine the means by which it occurs? The internalist explains this in terms of a non-backwards causation by representations of the end; but an externalist like Stout needs another account of how teleological explanation could make sense without backwards causation.

Larry Wright's account of teleological explanation gives the beginning of such an account. Roughly, Wright's account has it that x happens for the sake of y iff (i) x tends to bring about y , and (ii) x occurs because it tends to bring about y . So x is teleologically explainable in terms of y iff (i) and (ii) hold.

It may be useful to see how this works in the case of intentional action. Suppose we give an externalist teleological explanation of action, e.g. that Bob went to the refrigerator to get some beer. For this to be a good teleological explanation, according to Wright, it must be the case that goings to the refrigerator tend to bring about beer gettings, and that the event of Bob going to the refrigerator happened because of this. This looks like a reasonably intuitive picture of what is going on in this case.

Stout's account is not the same as Wright's; it's not easy, in fact, to see exactly what Stout's account is. But a key discussion is his account of what a means to an end is. The idea seems to be this: if we can give a noncircular account of what it is for one

event/process to a means to the end of another event/process, then we can say more, in terms of this means-end relation, about what teleological explanation is.

For some event m to be a means to some end e , two things must be the case: (i) there must be some way for e to occur which involves m , and (ii) m must be an available activity.

Circularity worries about ‘available activity.’

The worry that this definitions of ‘means’ and ‘end’ is too loose to do any real work.

2.2.2 Teleological explanation and intentional action

The teleological theory of action holds that some events are intentional actions in virtue of their being teleologically explainable. But this raises a problem: aren’t lots of things besides intentional actions teleologically explainable? And doesn’t this mean that the teleological theory of action will go wrong in counting these things as intentional actions? This is the objection that Stout tries to defuse toward the end of Ch. 3. He discusses two possible counterexamples.

Evolutionary explanation

Evolutionary change is the standard example of something which is teleologically explainable, but it is clearly not a species of intentional action. So Stout needs to refine his notion of teleological explainability to rule this out.

Stout distinguishes between strong and weak teleological explanations. His claim is that intentional action should be analyzed in terms of explainability by a strong teleological explanation, whereas evolutionary explanations are only a matter of weak teleological explanation.

The discussion here (110-112) is obscure and difficult to understand. But the basic idea seems to be this. When we say (for example) that we have eyes because having eyes is a means to the end of locating food, the phrase ‘having eyes’ is ambiguous. In the first usage, it refers to a property of all members of a population; in the second usage, it refers to a property of a (particular?) member of the population. If taken at face value, this seems crazy. In evolutionary explanation, we do not explain our having a property in terms of the utility of some one ancestor of ours having the property. So this cannot be what Stout has in mind. But what else could he have in mind? He says “What we cannot say in a case of evolutionary explanation is that the occurrence of some type of activity is explained in terms of that very type of activity being a means to an end.” But isn’t this exactly what we do in evolutionary explanation?

In other places, as in the first full paragraph on 112, he seems to have in mind a type-token distinction. It is true that in cases of evolutionary explanation, we explain my having some feature (e.g., this instance of the property of having eyes) in terms of the utility of *other instances* of possession of this property. And this does seem to be different from explanations of ordinary intentional action, in which we explain my hand’s moving in terms of *this particular instance* of the property of hand-moving being a means to an end.

But then we are back to the problem of making sense of teleological explanation. Wright's definition, for example, is stated in terms of what tokens of some type tend to bring about. If we focus only on some instance of a property being a means to an end, we have to give this up. So at this point, if we follow what seems to be the best route to distinguishing the kind of teleological explanation active in intentional action and in evolutionary explanation, I feel a bit in the dark about how we are supposed to be understanding teleological explanation as a special case of causal explanation.

A possibility: maybe we bring in causal processes here. To say the end for which a certain action is being done gives some information about which causal process the action is a part of, which perhaps gives some information about why it has happened. Is this the kind of explanation that Stout thinks that teleological explanation is? If so, why doesn't he say so?

A further question: Is having a feature like eyes supposed to be a kind of explainable activity? If not, how is this discussion of teleological explainability supposed to go together with the definition of the means-end relation in the preceding section? After all, something has to be a means to an end in order for it to be teleologically explainable.

Machines

A second class of counterexamples comes from simple cases of 'artificial intelligence.' The simplest example is that a thermostat might, it seems do something *in order to maintain the temperature in the room at 20*. The worry is that this makes it seem as though the behavior of the thermostat is teleologically explainable in terms of the end of maintaining the temperature in the room, although the thermostat is clearly not undertaking any intentional actions.

The discussion which aims to refine the definition of a means-end explanation to rule these out is on pp. 112-122. It is clear that Stout thinks that thermostats, internal organs, etc. fail to be agents because they fail to be adaptable in a certain way. How is this spelled out?

3 Results

3.1 *Intentional action*

The theory which Stout presents at the outset of the final chapter is:

A ϕ s intentionally \equiv_{df} *A*'s ϕ ing is governed by a method of justification which is normative and embodies a means-end sensitivity.

Stout spends a lot of time discussing different 'methods of practical justification.' This is actually kind of a puzzling phrase – what is such a method? To understand this, I think that we have to go back to Stout's discussion of *methods of deriving a description of the world* (16-18). He says little about what such methods might be, but they seem to be something like theories which, when given input, have as output propositions about the

world. Beliefs can be thought of as justified or not relative to a given method of deriving a description of the world. This would be an example, I take it, of a method of theoretical justification. A method of practical justification would be one which, instead of outputting propositions about the world, would output actions which are to be undertaken. It is a normative method in that it says what ought to be done, not a method for generating predictions about what will be done.

Methods of practical justification are those methods of justification which are normative and embody a means-end sensitivity.

What is it for an action to *governed* by such a method of practical justification? Stout glosses the idea as follows:

“Although [the method of practical justification] describes the process, it must *govern* the process. The process must result in whatever the method recommends. If the method came up with a different result, the process would have to adapt to that, rather than vice versa.”

The idea here seems to be that there is a kind of asymmetric counterfactual dependence of the process on the method. If the method changes its output, the process changes with it; but a change in the process does not tend to bring with it a change in the method.

Objections to the theory of intentional action

1. Do we really have an answer to the problems from the first chapter to the problems about actions caused by false beliefs?
2. The idea of analyzing intentional action in terms of notions which are then analyzed in terms of counterfactuals.
3. I still do not see how actions done for no reason, or for their own sake, are to be handled.

A possible counterexample to Stout's theory

As I understand it, a method of practical justification is something like a theory that yields conclusions about what ought to be done, and for an action to be governed by the theory is for there to be a certain kind of counterfactual dependence between the two. If this is right, then the following seems to be at least part of what Stout is claiming, where ‘*T*’ is a variable over theories of the right kind, and ‘ \vdash ’ for ‘entails’:

$$A \text{ } \phi \text{ s intentionally } \equiv_{df} \exists T \forall \psi (T \vdash A \text{ is to } \phi \ \& \ (T \vdash A \text{ is to } \psi \ \Box \rightarrow A \ \psi \text{ s}))$$

But there is a worry that this theory counts every action as intentional. Suppose that I trip on the chair while crossing the room. To see whether this act was intentional, we ask: is there a normative theory which (i) entails that tripping is to be done, and (ii) is such that if it had entailed that something other than tripping was to be done, I would have done that other action. But it seems likely that there will always be some such theory. (Consider simple theories, like the one which says that I should trip if there's an object in front of me, and not trip otherwise; and suppose that I'm very clumsy.)

3.2 Belief

Stout presents the following theory of belief:

A believes $p \equiv_{df}$ some teleological process which produces (some bit of) A 's behavior is governed by a method of practical justification which works on the assumption that p is true.

This raises a number of questions:

- *What makes some processes teleological?* Those which have a means-end sensitivity, i.e. which are such that a change in the end (e.g., the end of the process being made unattainable) will entail a change in the means (e.g., one appropriate to another end being pursued).
- *What makes a teleological process governed by a method of practical justification?* Counterfactual dependence on a theory ('method of justification') which yields normative conclusions.
- *What makes a method of practical justification work on the assumption that a certain proposition is true?* Here I think that we have to take the analogy between methods of justifications and theories fairly seriously. The idea, I think, is that assumption by a method of justification is something like use of a proposition in a theory in deriving a conclusion.

Objections to the theory of belief

1. The above counterexample to the theory of intentional action could be handled by requiring implicit knowledge of the theory in question. But that would make the account of belief circular.
2. How is the theory supposed to handle seemingly language-dependent beliefs?

3.3 Intention

Stout's theory of intention is similar to the theory of belief:

A intends to $\phi \equiv_{df}$ some teleological process which produces (some bit of) A 's behavior is governed by a method of practical justification which works on the assumption that ϕ ing is to be done.

This raises the same questions as above, plus one:

- *What makes a method of practical justification work on the assumption that a certain action is to be done?* Here the answer is presumably analogous to the one given above, about working on the assumption that a certain proposition is true.

Objections to the theory of intention

1. As Stout says, this makes intending to ϕ the same thing as believing that one should ϕ . But this seems wrong. One can clearly believe that one ought to perform some future action without intending to perform that action. (In my view, this is also a problem with any view of intentions as all-out evaluative judgements.)
2. How are desires and other pro-attitudes related to intentions, on this view?