Desires, Dispositions and Deviant Causal Chains

JOHN HYMAN

Abstract
Recent work on dispositions offers a new solution to the long-running dispute about whether explanations of intentional action are causal explanations. The dispute seemed intractable because of a lack of percipience about dispositions and a commitment to Humean orthodoxies about causation on both sides.

1.

The purpose of this article is to clarify the relationship between intentional action and desire, not as part of a full-dress analysis of either concept, but purely to shed light on the long-running dispute about whether explanations of intentional action are causal explanations. I shall argue that the dispute has seemed intractable because of a commitment to Humean orthodoxies about causation on both sides.

The dispute is not about every kind of explanation of intentional action. For example, we might explain why Henry bit a policeman by saying that he (Henry) was drunk. The act was almost certainly intentional. As a London magistrate pointed out in 1961, you don't bite policemen without intending it.[1] But the explanation doesn't indicate a motive. For all we have said, it might have been an affectionate nibble or a violent assault.

But while there are exceptions, *most* explanations of intentional action *do* indicate a motive, either explicitly or implicitly. This includes ones that mention desires or intentions ('Henry bit a policeman because he wanted to impress Lucy', '… in order to impress Lucy', '… with the aim or intention of impressing Lucy') and ones that mention reasons or beliefs ('… because it would impress Lucy', '… because he knew it would impress Lucy', '… because he believed it would impress Lucy').

[1] *Sunday Times*, 28 May 1961: *Sayings of the Week*, quoted in David Daube, *The Deed and the Doer in the Bible* (West Conshohocken, PA: Templeton, 2008), 34.

doi:10.1017/S0031819113000685 © The Royal Institute of Philosophy, 2013
First published online 10 September 2013

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

These are the explanations whose character is controversial. I shall concentrate for the most part on ones that mention intentions or desires, but my conclusions are meant to apply equally to ones that mention reasons or beliefs.

2.

Aristotle's statement in the *Nicomachean Ethics* is a convenient point of departure:

> The origin of action – its efficient, not its final cause – is choice, and that of choice is desire and reasoning with a view to an end.[2]

Not all intentional action, let alone all action, issues from instrumental reasoning or involves conscious choice. I take my keys out of my pocket when I get home with the intention of opening my front door, but I do not make a conscious choice to do so or rehearse the reason in my mind. But intentional action always (or almost always) involves desire, as long as 'desire' is understood in the broad sense it is often given by philosophers, so that it refers to the whole gamut of wanting and valuing, from the most egoistic to the most altruistic, from the most idealistic to the most mundane, and from the desire to live to a hundred, which depends on having abstract concepts, so that only language-users can have it, and whose satisfaction, if it occurs, may be very remote in time, to the appetites we share with non-human animals, such as hunger and thirst.[3]

All (or almost all) intentional acts involve desire in this broad sense, because an intentional act is always (or almost always) done either because the agent wants to do it or values doing it for its own sake, or in order to achieve something else she wants or values. If urges and cravings are not counted as desires, there will be some exceptions to this rule. For example, arguably, the compulsive hand-washer has no desire to wash her hands the twentieth time she does so on the same day, but finds it difficult to control the urge. But perhaps urges and cravings should not be distinguished from desires, in which case the hand-washer has an intense desire to wash her hands, which she may also want to resist.

[2] *Nicomachean Ethics*, 1139a.

[3] It is controversial whether one can do something, say, take one's keys out of one's pocket, *intentionally* or *with the intention of opening the door* without also intending to take one's keys out of one's pocket. The argument in this article is consistent with both views.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

I shall assume that a simple and conventional conception of desire is approximately right.[4] A desire is a disposition, in the broad sense of the term that corresponds to Aristotle's *hexis*. But desires differ in important ways from the simple physical dispositions that philosophers tend to focus on, such as fragility and solubility. The basic difference is that a simple physical disposition is manifested by causing or undergoing change, whereas a desire is manifested in two main ways: first, by purposive or goal-directed behaviour, specifically, behaviour aimed at satisfying the desire, in other words, at getting what it is a desire to have, or at doing what it is a desire to do; and second, by feeling glad, pleased or relieved if the desire is satisfied, and sorry, displeased or disappointed if it is frustrated.[5]

There can be other signs or symptoms of desire, but signs and symptoms are not manifestations. As Wittgenstein put it, they are not *criteria*.[6] *Only the manifestation of a desire is constitutive of its identity, and how it is defined*. This is true of dispositions in general. The green colour of a leaf is a reliable sign that it is photosynthetic, but only photosynthesizing is a manifestation of the disposition. Similarly, salivating can be a *sign* or *indication* of the desire to eat a steak, but eating a steak, and action aimed at getting and eating a steak, are not merely signs of the desire, they are manifestations.

The fact that a desire is manifested in goal-directed action has two direct consequences, both of which distinguish desires from simple physical dispositions, and even from primitive adaptive behavioural dispositions like the phototropic disposition of a plant. First, the object of a desire – i.e. what it is a desire for or a desire to do – and its manifestation in action can be related teleologically, as end and means. For example, Lucy's desire to dance can be manifested in dancing, but it can also be manifested in buying a tutu, as long as buying a tutu seems to Lucy conducive to dancing. Second, the

[4] A dispositional conception of desires differing in significant ways from the one set out here is proposed in Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994), 4.6.

[5] Empiricists from Locke to Mill insist on the intrinsic relation between desire, pleasure and pain. For example, Mill writes, 'desiring a thing and finding it pleasant, aversion to it and thinking of it as painful, are phenomena entirely inseparable or, rather, two parts of the phenomenon.' (J.S. Mill, *Utilitarianism* (New York: Macmillan, 1957), 49.)

[6] Wittgenstein distinguishes between symptoms and criteria in *The Blue and Brown Books* (Oxford: Basil Blackwell, 1958), 24*f*, and in *Philosophical Investigations*, third edition, trans. G.E.M. Anscombe (Oxford: Basil Blackwell, 1958), §354.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

way in which a desire is manifested in action depends on cognition: it depends on what the agent senses, knows and believes – for example, James's desire to please his mother will not be expressed by his going to church unless he knows or believes that doing so might please her. And it can be manifested in action that will actually displease her, if he is under a misapprehension about what she would like.

There is a good deal to be said for this simple conception of desire. First, it is clear that if one is disposed to do *x*, it does not follow that one wants to do *x*. One may be disposed to catch colds or to stammer when one is nervous without wanting to in the least. But it is equally clear that if one wants to do *x* one is disposed to do *x*, or something conducive to doing *x*. Someone who wants to do something, say dance, may be prevented or dissuaded from doing it, may want to do something else more, may feel lazy, or may change her mind. But as long as she wants to dance she is disposed to dance. The same applies to someone who wants to do something idiotic, such as punch a slot machine, and controls the desire so well that he never gets close to doing it. This is part of what distinguishes a desire from an idle wish. It also distinguishes a desire from a felt need. For although some desires, especially the appetites we share with other animals, are for things we need, it is conceivable, as Kenny points out, 'that a felt need might be like a felt ache: it might just cause one to hug oneself and lie immobile.'[7]

Second, it explains why one cannot want to get something one knows one already has – although of course one can want to keep it – and why wanting shades into hoping or wishing when one knows that one cannot do anything to satisfy one's desire, either because of the limitation of one's powers, such as hoping the weather will be fine next week, or because the thing is settled or in the past, such as wishing the weather had been fine last week.[8] (But the boundaries between wanting, hoping and wishing are not sharp. Our mental concepts draw distinctions in sfumato, so that one feature blends into another.)

Third, it explains the behaviourist idea, which Russell learned from the psychologist John Watson, that 'desire must be capable of being exhibited in action'.[9] But what exactly does this mean? In fact it is similar to Wittgenstein's remark that 'an "inner process"

[7] Anthony Kenny, *The Metaphysics of Mind* (Oxford: OUP, 1989), 34.

[8] Cf. Timothy Schroeder, *Three Faces of Desire* (Oxford: OUP, 2004), 16*ff*; G.F. Schueler, *Desire* (Cambridge, Mass.: MIT, 1995), 14.

[9] Bertrand Russell, *The Analysis of Mind* (London: George Allen & Unwin, 1921), 62.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

stands in need of outward criteria'.[10] It does not mean that one cannot desire something if one is unable to express the desire in action, so that an animal cannot be hungry if it is unable to seek food, say, because it is too weak. It means that if there is such a thing as the desire for X, then there is such a thing as exhibiting the desire for X in action. For example, there is such a thing as wanting to die, but there is no such thing as wanting, as opposed to wishing, one had not been born. And it means that one can have conflicting desires, such as the desire to eat trifle and the desire to be slim, and even contradictory desires, such as the desire to eat trifle and the desire not to eat trifle, but not a desire with a contradictory content, such as the desire to eat trifle and not to eat it. Wanting to have one's cake and eat it is not one desire, but two.

Fourth, it also explains Wittgenstein's criticism of Russell's theory of desire. Russell argued that the object of a desire, what it is a desire for, or to do, is the result that brings action manifesting the desire to an end, it is 'the state that brings quiescence':

> A hungry animal is restless until it finds food; then it becomes quiescent. The thing which will bring a restless condition to an end is said to be what is desired.[11]

Wittgenstein's well-known comment on the theory is this:

> I believe Russell's theory comes to the following: if I give someone an order and what he thereupon does pleases me, then he has carried out the order. (If I should like to eat an apple and someone gives me a punch in the stomach, so that the desire to eat goes away, then it was this punch that I originally wanted.)[12]

The objection is decisive. Russell's theory cannot be right because something that eliminates a desire is not necessarily what the desire is for. The more a desire is like a craving or an urge, the more plausible the theory seems. For example, an alcoholic in a play by Tennessee Williams – Brick, in *Cat on a Hot Tin Roof* – says that he drinks for the click, 'the click I get in my head when I've had enough of this stuff to make me peaceful'. In this case, Russell's remark that 'the thing which will bring a restless condition to an end is said to be what is desired' is a close fit. But most desires are not merely for

[10] *Philosophical Investigations*, §580.
[11] *The Analysis of Mind*, 32.
[12] Ludwig Wittgenstein, *Philosophical Remarks*, (ed.) R. Rhees (Oxford: Basil Blackwell, 1964), 64.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

quiescence, and Wittgenstein's example shows that unless that *is* what a desire is for, eliminating it and satisfying will not necessarily be the same thing. Like any disposition, the *object* of a desire – what it is a desire *for* or *to do* – is constitutive of its identity and how it is defined, whereas what eliminates it is not.

Finally, the simple conception of desire proposed here explains why the object of a desire, like the object of any disposition, is standardly identified in English with an infinitival clause, and not, as philosophers who believe desires are propositional attitude would prefer, a sentential clause.[13] It is true that the object of a desire is also commonly identified with a nominal clause, but if one desires a thing, such as an apple or a house, one desires to eat it, occupy it or inhabit it, or minimally have it under one's control.

This simple conception of desire may not provide an accurate characterization of every desire or a complete characterization of any desire. Perhaps some desires are too fleeting to be dispositions, although perhaps dispositions can be fleeting too. And desires can be manifested in ways I have not discussed, for example, in obsessions, preoccupations, and subtler patterns of thought. But it is sufficient for our present purposes if it provides a partial characterization of many desires, and it certainly does that. As I have said, I shall assume that it is approximately right.

Now if desires are dispositions, then presumably they explain action in the way dispositions in general explain action. But it is sometimes said that explanations that refer to dispositions are vacuous or uninformative, and explain nothing at all. The complaint has some merit when the disposition mentioned is to produce the effect whose cause we want to know. Moliére's joke about a student who is congratulated for explaining that taking opium makes one fall asleep because of its *virtus dormitiva* – i.e. its soporific power – is a case in point.[14] The explanation is not quite vacuous, since it excludes the hypothesis that opium activates the *virtus dormitiva* of some other popular part of the French diet, such as asparagus or red wine, but its value is limited. It is like explaining that Henry kicked a lamp-post because he wanted to, which excludes the hypothesis that he did it under duress, but is likely to leave us wanting to know more.

But explanations that refer to dispositions are not all alike. For example, 'Lead is poisonous because it is a neurotoxin' is a better

[13] There is an astute discussion of the idea that beliefs and desires are propositional attitudes with opposed 'directions of fit' in Maria Alvarez, *Kinds of Reasons* (Oxford: OUP, 2010), 66–71.

[14] *Le Malade Imaginaire*, III.iii.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

explanation than the one Molière lampooned, because it excludes a larger range of alternatives. Most explanations referring to desires are better than 'Because he wanted to' for the same reason. All explanations without exception are informative to the extent that they rule out alternatives, and since one act can be expected to have many different consequences it can express many different desires. For example, one can open a window because one wants to let the cool air in, because one wants to let the smell of frying bacon out, because one wants to hear what the neighbours are saying, and so on. The right explanation will say which desire or desires the act actually manifested, and rule out the rest.

3.

The idea that desires cause the acts we do with the intention of satisfying them was never disputed before the twentieth century, except by those, such as Malebranche, who denied the existence of psycho-physical and physical causation entirely. The controversy was initiated by Wittgenstein's writings in the 1930s, and the arguments he and his followers advanced rely on different elements of what I shall call, with deliberate vagueness, the Humean (as opposed to Hume's) theory of causation. I use the term to refer indiscriminately to various combinations of doctrines or assumptions about causation which were defended by or have been imputed to or derive from Hume, and not to refer to Hume's own theory of causation in particular.

These are the principal arguments:

- ***Causes are events, which are distinct from and precede their effects***; whereas desires are not events, and in most cases, if one says what desire an act was meant to satisfy, one does not identify a feeling, image or idea that precedes the act the desire explains: one does not answer the question 'what did you see or hear or feel, or what ideas or images cropped up in your mind and led up to it?'[15]

- ***Our knowledge of causes is uncertain, and relies on inductive inference from a plurality of cases***, because ***causation cannot be observed in any particular case***; whereas our knowledge of our own motives is normally certain and non-inferential.[16]

[15] G.E.M. Anscombe, *Intention* (Oxford: Blackwell, 1957), 17.
[16] Wittgenstein, *The Blue and Brown Books*, 15.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

- ***A singular causal statement is implicitly general: it implies that events resembling the cause are invariably followed by events resembling the effect***. But the statement that a particular act was done with a particular motive does not imply that the same motive invariably leads to the same act. For instance, a defendant's assertion that he did a certain act because of threats does not imply that he would act in the same way if he was threatened again, or that he or others always respond to threats in that way.[17]

- ***It is conceivable that any kind of cause should have had a different kind of effect from the one it actually has***, whereas what a desire is *for*, or *to do*, is constitutive of its identity and how it is defined. For example, it is not a contingent fact that a desire for an apple is satisfied by an apple rather than a punch in the stomach, or a pear.[18]

Of course, none of these arguments has to be taken to show that desires are not causes, even if it is accepted that it shows that desires are not causes *as the Humean theory conceives of them*. It can be taken to reveal a defect in the Humean theory instead. And if it *is* taken this way, different diagnoses regarding the defect in the theory are possible.

Davidson's ingenious and influential response to these arguments is minimalist: a refinement of the Humean theory – perhaps it is no more than a clarification – is sufficient to meet them.

- True, desires are not events; but the 'onslaught' of a desire is, such as when the smell of food makes one feel hungry; and so is a perception that triggers the manifestation of a desire, such as when catching sight of its mother makes a child run towards her.
  (So desires are better described as causal factors than as causes; but explanations that refer to desires *are* causal explanations.)

- True, our knowledge of our own motives is normally certain and non-inferential; but Hume himself points out that we can 'attain knowledge of a particular cause merely by one experiment.'
  (In fact Hume adds the qualification, 'provided it be made with judgment and after a careful removal of all foreign and

[17] H.L.A. Hart & A.M. Honoré, *Causation in the Law*, second edition (Oxford: OUP, 1985), 55*f*.

[18] A.I. Melden, *Free Action* (London: Routledge & Kegan Paul, 1961), 52*ff*.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

superfluous circumstances.'[19] But we normally know what desire motivated us in the ordinary circumstances of our daily lives, and not just in an experimental situation that is carefully contrived to exclude confusion about motives.)

- True, the statement that an act was done because of a certain motive does not imply that the same motive would always lead to the same act; but it can imply that a law covering the events concerned exists without implying that the law can be stated in the language of motives and desires. If desires are physical states it can be stated in the language of physics.

- True, what a desire is for is constitutive of its identity *qua* desire; but it does not follow that it is constitutive of its identity *qua* physical state.

I agree with Davidson that desires are causal factors, but I think the arguments of Wittgenstein and his followers exposed a more serious flaw in the Humean theory of causation. The root of the problem, the reason why it seemed to them that desires could not be causal factors, is that the Humean theory cannot explain the exercise of dispositions. But neither side in the dispute saw this. (Of course Hume himself never intended to explain the exercise of powers, but only to explain how the *idea* of power can be produced by impressions of sequences of events.) In effect – although I do not mean to suggest that this is how it appeared – Wittgenstein and his followers faced a choice between challenging the Humean theory and excluding causation by desires, and they made the wrong choice. For his part, Davidson adhered just as closely to the Humean theory, and the principal difficulty faced by his solution – eliminating the 'deviant' causal chains from desires to acts – which Anscombe and others sympathetic to Wittgenstein's position seized on, is a residual sign of the same weakness in the Humean theory.

In sum, the opposing views about explanations of intentional action arrived at an impasse because of a failure of percipience on both sides regarding dispositions. But it can be broken if this is put right. Or so I shall argue.

If this account of the matter is correct, the four arguments above should be answered rather differently. Expressed with the same brevity, the answers should be roughly as follows.

[19] David Hume, *A Treatise of Human Nature*, (ed.) D.F. Norton & M.J. Norton (Oxford: OUP, 2000), 1.3.8.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

- True, desires are not events; they are dispositions. But dispositions are causal factors, which belong to the causal history of the acts and events they explain.

- True, our knowledge of our own motives is normally certain and non-inferential. But there is no inherent difficulty in the idea that knowledge of a singular causal fact can be certain and non-inferential, unless we accept the Humean doctrine that a singular causal statement is made true by a regularity in the general pattern of events, rather than by anything present locally in the particular case.

- True, the statement that an act was done because of a certain motive does not imply that the same motive would always lead to the same act; but it is normal for many dispositions to contribute to a particular act, and when this happens, the outcome depends on the exact mix.

- True, what a desire is *for*, or *to do*, is constitutive of its identity, and how it is defined; but this is true of every disposition.

I said that the opposing views about explanations of intentional action arrived at an impasse. Here it is. On the one hand,

> (A) if James went to church *and* wanted to please his mother (and knew or believed that going to church would please her), it does not follow that he went to church *because* he wanted to please his mother.

On the other hand,

> (B) if James went to church *and* wanted to please his mother (and knew or believed that going to church would please her), and his churchgoing was caused by his desire to please his mother, it *still* does not follow that he went to church *because* he wanted to please his mother (as this would normally be meant and understood), because the 'causal chain' running from the desire to the act could be 'deviant' or 'freakish'.

(A) suggests that explanations of intentional action are causal explanations. Davidson writes:

> If, as Melden claims, causal explanations are 'wholly irrelevant to the understanding we seek' of human action then we are without an analysis of the 'because' in 'He did it because …', where we go on to name a reason. Hampshire remarks, of the relation between

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

> reasons and action, 'In philosophy one ought surely to find this … connection altogether mysterious'. Hampshire rejects Aristotle's attempt to solve the mystery by introducing the concept of wanting as a causal factor […] But I would urge that, failing a satisfactory alternative, the best argument for a scheme like Aristotle's is that it alone promises to give an account of the 'mysterious connection' between reasons and actions.[20]

Arguably, Davidson's Humean conception of efficient causes is so different from Aristotle's that it was misleading of him to present himself as a champion of Aristotle's doctrine at all. Be that as it may, (B) shows that postulating a causal connection between the desire and the act does not provide 'an analysis of the "because"', or explain the '"mysterious connection" between reasons and actions'. Anscombe writes:

> [Davidson's solution] lacks acumen. True, not only must I have a reason, it must also "operate as my reason": that is, what I do must be done *in pursuit* of that end and *on grounds* of the belief. But not just *any* act of mine caused by my having a certain desire is done in pursuit of the object of desire; not just *any* act caused by my having a belief is done on grounds of the belief. Davidson indeed realises that even identity of description of *act done* with *act specified in the belief*, together with causality by the belief and desire, isn't enough to guarantee the act's being done *in pursuit of the end* and *on grounds of the belief*. He speaks of the possibility of "wrong" or "freak" causal connexions. I say that any recognisable causal connexions would be "wrong", and that he can do no more than postulate a "right" causal connexion in the happy security that none such can be found. If a causal connexion were found we could always still ask "But was the act done for the sake of the end and in view of the thing believed?"[21]

This is not perfectly clear: 'any recognisable causal connexions would be "wrong"' suggests that *no* act caused by a desire is done with the intention of satisfying it whereas 'not just *any* act of mine caused by my having a certain desire …' suggests, on the contrary, that some intentional acts *are* caused by desires, but that a causal connection between a desire and an act does not guarantees that the act is done with the intention of

[20] Donald Davidson, 'Actions, Reasons, and Causes', in *Essays on Actions and Events* (Oxford: OUP, 1980), 11.

[21] G.E.M. Anscombe, 'Von Wright on Practical Inference', in P.A. Schilpp & L.E. Hahn (eds), *The Philosophy Georg Henrik von Wright* (La Salle: Open Court, 1989), 378.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

satisfying it. Be that as it may, the principal point is that Davidson 'can do no more than postulate a "right" causal connexion in the happy security that none such can be found.' And by the time Anscombe had published this remark, Davidson had decided that this was true. 'Several clever philosophers have tried to show how to eliminate the deviant causal chains,' he wrote, 'but I remain convinced that the concepts of event, cause, and intention are inadequate to account for intentional action.'[22]

If this is true, it does not follow that when wants explain intentional acts they are *not* causal factors, or that 'any recognisable causal connexions would be "wrong"'. But it *does* follow that if the relation between intentional acts and desires is mysterious, the idea that desires are causal factors will not dispel the mystery. Hence the impasse.

In outline, the solution I propose is this.

First, the fact that there can be a 'deviant' causal connection between a desire and an act that is conducive to satisfying it has nothing particularly to do with reasons or desires. The reason is simply that desires are dispositions, and with certain exceptions, such as radioactive decay and rest mass, to which I shall return, every disposition can be connected to the kind of occurrence that normally manifests it by a 'deviant' or 'freakish' causal chain. Take Molière's example of a disposition, the *virtus dormitiva*. A man might take a soporific drug before driving, and the drowsiness induced by the drug might make him crash the car and knock himself unconscious. If this happened, he would lose consciousness because he took the drug, but the exercise of its *virtus dormitiva* would be preempted by the crash.

Second, if Anscombe and Davidson are right in thinking that the 'deviant' causal chains from a desire to an act cannot be eliminated – in other words, if the 'right' causal chain cannot be defined, except as one where the act is done 'in pursuit of that end' – again, this has nothing particularly to do with intentional action as such. It has nothing particularly to do with 'the normative character of psychological concepts' (*contra* Davidson), and it is not because teleological explanations are not causal explanations (*contra* Anscombe). It is because *the exercise of a disposition cannot be defined as a specific kind of concatenation of events*.[23]

[22] Donald Davidson, 'Problems in the Explanation of Action', repr. in *Problems of Rationality* (Oxford: OUP, 2004), 106.

[23] Davidson moves some way towards acknowledging this in 'Problems in the Explanation of Action'. See also Ralf Stoecker, 'Climbers, Pigs and Wiggled Ears: The Problem of Waywardness in Action Theory', in

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

Third, if the 'deviant' causal chains between taking a soporific drug and losing consciousness cannot be eliminated, it does not follow that the fact that a man took the drug explains why he lost consciousness *causally* if the causal chain was 'deviant' and *non-causally* if it was not. It obviously explains it causally in both cases. But the causal chain can be described as 'deviant' in the case of the man who crashes his car *because the disposition is not manifested*. Exactly the same is true of desires. It is a mistake to imagine that a climber's wanting to rid himself of the weight and danger of holding another man on a rope explains why he loosened his hold on the rope *causally* if the connection was 'deviant' and *non-causally* if it was not. His desire is a causal factor either way, but the process can be described as 'deviant' if letting go is an effect or 'symptom' but not a manifestation or 'criterion' of the desire.

Finally, many explanations that refer to the agent's own reasons for doing an act, and every explanation that refers to his aim or intention in doing it, contains non-redundant information about the desires manifested by the action. For example, 'James's reason for going to church was that it would please his mother' and 'James went to church with the aim (or intention) of pleasing his mother' contain the information that James's going to church manifested his desire to please his mother. Hence, if desires are causal factors, both of these explanations contain non-redundant information about causal factors, and this makes them causal explanations, in one perfectly good sense of the phrase.

I shall explain this in more detail in sections 4 and 5.

4.

Wittgenstein wrote about intentional action in every phase of his career, but his influence on the controversy about explanations of

---

S. Walter & D. Heckmann (eds), *Physicalism and Mental Causation* (Exeter: Imprint Academic, 2003). Others have argued that the idea of causation by dispositions should play a key role in solving the problem of deviant causation by desires (e.g. Kieran Setiya, *Reasons without Rationalism* (Princeton NJ: Princeton Univ. Press, 2007), 32), or pointed out that deviant causation by desires is a special case of deviant causation by dispositions in general (e.g. David Armstrong, 'Beliefs and Desires as Causes of Actions: A Reply to Donald Davidson', *Philosophical Papers* **4** (1975), 4*f*; Nomy Arpaly, *Merit, Meaning and Human Bondage* (Princeton NJ: Princeton Univ. Press, 2006), 47 & 71).

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

intentional action is due to his writings in the 1930s, especially the following passage in the *Blue Book*:

> The proposition that your action has such and such a cause, is a hypothesis. The hypothesis is well-founded if one has had a number of experiences which, roughly speaking, agree in showing that your action is the regular sequel of certain conditions which we then call causes of the action. In order to know the reason which you had for making a certain statement, for acting in a particular way, etc., no number of agreeing experiences is necessary, and the statement of your reason is not a hypothesis. The difference between the grammars of "reason" and "cause" is quite similar to that between the grammars of "motive" and "cause". Of the cause one can say that one can't *know* it but can only *conjecture* it. On the other hand one often says: "Surely *I* must know why I did it" talking of the *motive*. […] The double use of the word "why", asking for the cause and asking for the motive, together with the idea that we can know, and not only conjecture, our motives, gives rise to the confusion that a motive is a cause of which we are immediately aware, a cause 'seen from the inside', or a cause experienced.[24]

Wittgenstein refers here to motives rather than desires, but this is not an important difference, and in fact he made similar remarks about desires to Waismann, which are recorded both in Waismann's notes of their conversations and in *The Principles of Linguistic Philosophy*, which was based on dictations by Wittgenstein.[25]

Anscombe points out in *Intention* why this argument is unconvincing. She comments on a case where someone is startled by a face appearing at a window: 'the subject is able to give the cause of a feeling or bodily movement in the same kind of way as he is able to state the place of his pain or the position of his limbs.'[26] These kinds of statements – for example, 'The pain is in my ankle' or 'My legs are crossed' – do *not* seem to be conjectures or hypotheses, whose credibility depends on 'a number of agreeing experiences'. And Anscombe is surely right to suggest that saying what made one start is similar. We *are* immediately aware of causation in this case – 'a cause experienced' is

[24] *The Blue and Brown Books*, 15.

[25] See Gordon Baker, *The Voices of Wittgenstein: The Vienna Circle* (London: Routledge, 2003), 424 *ff* and F. Waismann, *The Principles of Linguistic Philosophy* (London: Macmillan, 1965), 123.

[26] *Intention*, 10.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

an apt description, just as it is when one is stung by a bee or knocked over by a blow – and there is no obvious reason why we should not also normally be able to state the desire that caused us to act, in a similar way. If singular causal statements were invariably conjectures, Wittgenstein would be right in saying that stating a motive is not stating a cause, because stating a motive is not normally a conjecture. Unconscious motives, such as the ones Freud postulated to explain verbal slips, are exceptions. But examples like the one Anscombe mentions disprove Wittgenstein's assumption.

Anscombe does not endorse the Humean doctrine that a particular instance of causation counts as such only by conforming to a general pattern, which Wittgenstein took for granted when he wrote the passage quoted above. On the contrary, she implies that our ability to give the cause of a feeling or bodily movement, such as when we are startled, is incompatible with it.[27] But she still insists, now in agreement with Wittgenstein, that explanations of intentional action are not generally causal explanations, because she assumes that causes are events.[28] The desire an act is intended to satisfy, she points out, is not generally a 'mental cause' in the sense in which catching sight of a face at the window is, because it is not generally 'an event that brings the effect about'.[29] There are exceptions, 'feelings of desire' such as pangs of hunger, but in most cases, if one says what desire an act was meant to satisfy, one does not answer the question 'what did you see or hear or feel, or what ideas or images cropped up in your mind and led up to it?'

But if the desire is not 'an event that brings the effect about', how does it explain the act? Her answer is that it *interprets* it.

> Motives may explain actions to us; but that is not to say that they "determine", in the sense of causing, actions. We do say: "His love of truth caused him to …" and similar things, and no doubt such expressions help us to think that a motive must be what produces or brings about the choice. But this means rather "He did this in that he loved the truth"; it interprets his action.[30]

[27] This is how I interpret her remark about Hume's theory of causation quoted below in §7. Her doubts about the Humean theory of causation bore fruit, more than a decade later, in G.E.M. Anscombe, 'Causality and Determination. An Inaugural Lecture', repr. in E. Sosa *Causation and Conditionals* (Oxford: OUP, 1975).

[28] Anscombe contrasts mental causes with reasons on p.10 of *Intention*, with motives on p.16, and with intentions on page 17.

[29] *Intention*, 17*f*.

[30] Ibid., 18*f* & 24.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

What kind of interpretation does Anscombe have in mind? If a doctor interprets a tremor as a symptom of a disease, the interpretation *is* a causal explanation. The text is elusive, and the phrase 'in that' is unhelpful, since it occurs in different kinds of explanations, like 'since' and 'because'. For example, 'Let him die, in that he is a fox' is a justification, whereas 'For in that he himself hath suffered being tempted, he is able to succour them that are tempted' is a causal explanation.[31] But although it is hard to be certain, she seems to believe that motives like the desire for gain (another of her examples) and the love of truth explain an act by identifying the good, or apparent good, the agent aimed to realize or achieve, in other words, what it was about the act that made it seem worth doing – say, that it was (so he thought) likely to enrich him, or that it served (so he thought) the interests of truth.

If this is an accurate summary, Anscombe denies that an explanation of an intentional act in terms of the agent's motive or desire is a causal explanation because of a combination of two reasons. First, with the exception of pangs of hunger, sudden impulses to kiss lovers and the like, desires are not events, and *a fortiori* not events that trigger actions. Second, explanations of intentional acts that refer to non-trigger-like desires interpret them, in the way explained. As she also puts it, they 'say something like "See the action in this light".'[32] If she is right, empiricists like Locke and Russell thought of all desires as if they were urges and cravings. They regarded the exceptional case as the normal one, and so they imagined that explanations of conscious purposive behaviour by human beings postulate efficient causes – uncomfortable feelings that intitiate behaviour – when in fact they postulate final causes, aims or goals towards which the behaviour is directed.

But neither reason Anscombe gives for doubting whether explanations referring to desires are causal explanations is convincing. Regarding the first, we have seen that causal explanations do not invariably refer to triggering events. Since the desire for gain and the love of truth are abiding properties, asking whether a man launched a Sunday paper out of one or the other is *not* like asking what kind of sensation made Brick pour himself a vodka, and cannot be made to look like one without changing its meaning. It is more like asking whether a ship displaced a certain quantity of water because of its size or its weight, or whether a flower attracted a bee because of its colour or its scent. But these are evidently causal questions

[31] *2 Henry VI*, III.i; *Hebrews* 18.2.
[32] *Intention*, 21.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

about the events concerned, and the answers are evidently causal explanations. So we have as yet no reason to deny that explanations which refer to desires are causal explanations too.

Regarding Anscombe's second reason, it is true that 'His love of truth caused him to …' and 'He did it to make money' identify the good or apparent good the agent wanted to realize or achieve. But we do not have to choose between saying this and saying that these explanations are causal, and we have every reason to say both. For *what* the man desires, in other words, his aim or goal – *to make money* or *to serve the truth* – is an intensional object, an object of thought, and therefore *not* a causal factor. But the desire itself is a disposition to act in ways that seem conducive to satisfying it, and so presumably *is* a causal factor. So the explanations *do* interpret the act, in the way Anscombe suggests; but they *also* explain it causally. They attribute a justification to the agent, because they say what kind of good he saw in doing the act; but they wrap up the justification in a causal explanation, because they say that seeing this kind of good in the act is what caused him to do it. Russell's assimilation of desires to urges and cravings was a mistake because it led him to equate the object of a desire with whatever extinguishes it, and thereby to eliminate teleology from the definition of desire, not because explanations of intentional action are not causal explanations.

If this is right, a desire is part of the causal history of the act or event it explains, just as a person's other physical or mental dispositions are: 'Cecilia came out in a rash because she is allergic to peanuts', 'Abe blushed because he is shy', 'Sam laughed because he was nervous', 'James went to church because he wanted to please his mother', and so on. Of course there are important differences between desires and simple physical dispositions, as we have seen, and there are also important differences between the desires of adults and older children and the desires of animals and infants, inasmuch as the former normally know what they desire; many of their desires can be assessed as reasonable or unreasonable, wise or foolish; and their desires normally leave room for choice, by which I mean that they are normally able to resist exhibiting them in intentional action. But for all that, explanations that refer to dispositions are *echt* causal explanations, whatever kind of dispositions they refer to. *How* they explain, exactly what part of a causal story they tell, and whether a disposition is the cause, or part of the cause, of its manifestation – these are contentious questions.[33] But *that* explanations

[33] For a detailed discussion of the causal relevance of dispositions, see J. MacKitrick, 'Are Dispositions Causally Relevant', *Synthese* **144** (2005).

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

that refer to dispositions are causal explanations should be beyond doubt.

5.

By common consent, the most difficult objection faced by the claim that desires are causal factors arises from the existence of deviant causal chains. The best-known example is Davidson's, which I mentioned earlier. Here it is again, in his own words:

> A climber might want to rid himself of the weight and danger of holding another man on a rope, and he might know that by loosening his hold on the rope he could rid himself of the weight and danger. This belief and want might so unnerve him as to cause him to loosen his hold, and yet it might be the case that he never *chose* to loosen his hold, nor did he do it intentionally.[34]

Davidson and Anscombe agree that we cannot distinguish the case where the climber loosens his hold because he wants to rid himself of the weight and danger of holding another man on the rope from the case where his desire causes him to loosen his hold 'deviantly' by defining the 'normal' causal route from the desire (or its 'onslaught', or an event that triggers it) to the act. But this is not a problem about desires in particular, it is about dispositions and powers in general.

Deviant causal chains have been around in philosophy for longer than we might think. In fact, occasionalism is the arch-theory of deviant causal chains.[35] According to Malebranche, the pre-eminent early-modern occasionalist, there is no psycho-physical or physical interaction, but God's will ensures that whenever a physical or mental event occurs, the events we are inclined to think of as their effects follow. For example, when a person walks or speaks, we tend to attribute the motion in her legs or her lips to *her* agency, but in fact

[34] Donald Davidson,'Freedom to Act', repr. in *Essays on Actions and Events*, at page 79.

[35] On the history of occasionalism, see M. Fakhry, *Islamic Occasionalism* (London: George Allen & Unwin, 1958); D. Clarke, 'Causal Powers and Occasionalism from Descartes to Malebranche', in S. Gaukroger (eds), *Descartes' Natural Philosophy* (London: Routledge, 2000); S. Nadler, 'Malebranche on Causation', in S. Nadler (ed.), *The Cambridge Companion to Malebranche*, ed. S. Nadler (Cambridge: CUP, 2000).

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

the only power that is manifested is God's power, and so the only real agency is God's agency. She has a volition to move her body in a certain way, but God causes the motion. Exactly the same is true when she pricks her finger and feels pain, and also when one billiard ball strikes another one and the second ball begins to move. It seems to us as if the pin causes the pain and the first ball sets the second ball in motion, but in reality it is God. In Malebranche's system, God's monopoly on power is absolute, and *every* sequence of events that counts as cause and effect by Humean criteria is 'deviant', except for the ones where the cause is the will of God.

Hume has been described as Malebranche without God, because he reduces causation to the recurring sequences of events that remain when God's agency is removed from Malebranche's system. But if we start out with occasionalism and eliminate God's power to cause events, we eliminate the distinction between deviant and normal causal chains as well. For if causation consists simply in the regular or law-like concatenation of events, one causal chain may be more circuitous or more suprising than another, but the question 'Deviant or normal?' cannot arise, because there is no power in the world from which an act or event can either proceed, if the chain is normal, or fail to proceed, if it is not. This is why the Humean theory of causation cannot 'eliminate the deviant causal chains', for example, why it cannot distinguish between the case where the climber loosens his hold *in order to* kill the other man and the case where his desire to kill the other man causes him to loosen his hold without him doing so for this reason. This is just a particular instance of a more general failure to distinguish between 'true' causes and 'occasional' causes. We deem the causal chain to be normal if, and only if, the climber loosens his hold *in order to* kill the other man; and in general we deem the process running from a power's standard trigger to an act or event of the kind that manifests it to be normal if, and only if, this act or event is in fact a manifestation of the power. But this process cannot be reduced to a specific kind of concatenation of events.

The reduction of powers in modern philosophy was originally part of the campaign to liberate science from Aristotelianism, which was initiated by Galileo and Descartes. But since powers cannot be perceived with the senses, there is a perennial tendency for empiricists to clarify, or falsify, or reform our talk about them (which verb seems apt depends on one's point of view).

In the twentieth century, most reductionists about powers employed the method of contextual definition. Instead of identifying powers with geometrical structures, as Descartes had done, or defending Hume's notorious claim that the distinction between a

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

power and its exercise is 'entirely frivolous',[36] reductionists argued that categorical statements referring to powers can be explicated as hypothetical statements relating the stimulus that triggers the power to the response that manifests it. Armstrong summarizes the approach when he claims that powers are 'congealed hypothetical facts or states of affairs'.[37] But in recent years the reductionist programme has come under sustained attack. I shall not try to prove that its opponents are right. It is enough, for my purposes, to show that the question of whether the deviant causal chains between a desire and an act can be eliminated can be subsumed under the same general question about powers.

The simplest approach to contextual definition is to hold that the statement that an object or substance has a certain power is revealed by analysis to be the hypothetical statement that if the appropriate stimulus occurs the response follows. For example, Ryle writes as follows:

> Dispositional statements are neither reports of observed or observable states of affairs, nor yet reports of unobserved or unobservable states of affairs. […] To say that this lump of sugar is soluble is to say that it would dissolve, if submerged anywhere, at any time and in any parcel of water.[38]

Generalized, and relativized to a specific time, Ryle's analysis can be expressed thus:

> (C) A substance *x* is disposed at time *t* to give response *r* to stimulus *s* if, and only if, *x* would give response *r* if *x* were to receive stimulus *s* at time *t*.

This is known as the conditional analysis of dispositions, or sometimes the 'simple' or 'naïve' conditional analysis. As these names suggest, more complex and sophisticated conditional analyses have been proposed, which are designed to avoid some of the counterexamples to (C), but I shall confine my remarks to (C), and direct readers to more comprehensive studies in the notes.[39]

[36] *A Treatise of Human Nature*, 1.3.14.

[37] D.M. Armstrong, *A World of States of Affairs* (Cambridge: CUP, 1997), 79.

[38] Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949), 107*f*.

[39] See especially George Molnar, *Powers* (Oxford: OUP, 2003), ch.4; Alexander Bird, *Nature's Metaphysics* (Oxford: OUP, 2007), ch.2. Troy Cross, 'Recent Work on Dispositions', *Analysis* **72** (2012) is a good survey.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

(C) is vulnerable to several kinds of counterexample.

To begin with, George Molnar argues that some dispositions are exercised without a stimulus or trigger.

Evidently, most action is interaction, and so (i) *most dispositions are powers to interact*. This includes the examples that crop up in the literature most often, such as the solubility of sugar in water and the brittleness of glass. And in most cases, (ii) *the exercise of a complementary pair of active and passive dispositions is triggered by the agent and the patient coming into contact in a certain way*. Sugar is submerged in water; the brittle sheet of glass is struck. But there are exceptions to both generalizations. Regarding (i), an unstable atom's propensity to decay is not exercised in any kind of interaction, it simply disintegrates in a specific way. Similarly, some desires are to interact with apples, lovers, and so on, in various ways, while others, such as the desire to die or sleep or compose a sonnet, are not. Regarding (ii), radioactive decay occurs spontaneously, without any kind of trigger, and a body's rest mass – according to General Relativity, its disposition to deform space-time – does not have a trigger either, not in this case because it is exercised spontaneously at a specific time, but because it is exercised continuously, for as long as the body exists. As Molnar puts it, it does not have a toggle, it cannot be switched on or off.[40]

Does the expression of desire in action invariably have a trigger? It is not *obvious* that it does, but it is plausible. I may often have no reason – i.e. no justification – for doing something at a particular moment, rather than a few moments earlier or later, but it does not seem plausible that it can be a matter of pure chance. In some cases, the trigger is contact, but in more cases, it is seeing, hearing or smelling something – contact at a distance, as it were – as when a child catches sight of its mother and runs towards her. And in many cases the trigger is inside the agent's body and we do not know what it is. For example, I want to tell the children to be quiet, but say nothing, hoping they will settle by themselves. After a while, I say something. Perhaps the volume of chatter rose above a certain threshold, perhaps my blood sugar dropped below a certain threshold. I may not know why I spoke exactly when I did. Whatever the trigger was on this occasion, it was not necessarily a thought or feeling or perception.

If radioactive decay and rest mass are dispositions without triggers, then (C) cannot provide a template for the contextual definition of every disposition, and Armstrong's claim that powers are 'congealed

[40] *Powers*, 86. Both examples are due to Molnar.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

hypothetical facts or states of affairs' cannot be true without exception. But I do not think the reductionist should be troubled by this conclusion, because a body's rest mass and an unstable atom's propensity to decay are reducible to strict – that is, indefeasible – laws, deterministic in the first case and stochastic in the second. Perhaps it is a mistake to regard them as dispositions at all. As Strawson says, 'in the most sophisticated reaches of physical theory […] causation is swallowed up in mathematics.'[41] What *should* trouble the reductionist is that there are various kinds of counter-examples to (C) involving dispositions that *do* have triggers, which suggests that *only* dispositions without triggers are reducible to strict laws.

Some of these counter-examples rely on the fact that the process running from a disposition's trigger to its manifestation takes some time, so an intervention can prevent the manifestation from occurring. So if a device is set up to intervene in this way if the stimulus occurs, it will not be true that the stimulus would cause the manifestation, or that if the stimulus occurred the manifestation would follow. One well-known example of this kind is due to C.B. Martin.[42] A wire is live if, and only if, it has the disposition to transmit a current when it is touched by an earthed conductor; it is dead if, and only if, it is not live. But suppose a live wire is attached to a safety-device which detects when the wire is touched, and makes it dead before the current flows. If the wire is touched at t it will be live but it will not transmit a current, and so the right-hand side of (C) is false. The device could also have a 'reverse cycle', so that it detects when a dead wire is touched and makes it live. In this case, the right-hand side of (C) will be true although the wire is dead.

Mark Johnston describes a different kind of example:

> A gold chalice is not fragile but an angel has taken a dislike to it because its garishness borders on sacrilege and so has decided to shatter it when it is dropped. Even though the gold chalice would shatter when dropped, this does not make it fragile because […] something extrinsic to the chalice is the cause of the breaking.[43]

Or, leaving the realm of fantasy, something that is not fragile, such as a sheet of toughened glass, could have an explosive device with a sensitive detonator attached to it, so that it would be shattered by the

[41] P.F. Strawson, *Analysis and Metaphysics* (Oxford: OUP, 1992), 120.

[42] See C.B. Martin, 'Dispositions and Conditionals', *Philosophical Quarterly* **44** (1994).

[43] Mark Johnston, 'How to Speak of the Colours', *Philosophical Studies* **68** (1992), 232.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

explosion if it were dropped.[44] In this case, the hypothetical statements are true, but the glass is not fragile. This example does not rely on the possibility of switching the disposition on or off after the trigger has occurred. Instead, like the example of the driver who takes a soporific drug and crashes his car, it relies on the fact that a response will only count as a manifestation of a disposition if it is the result of the right kind of process. As Molnar puts it, the relation between a disposition and its manifestation is *process-specific*, at least to some degree, whereas the relation between a cause-event and an effect-event is *process-unspecific*.[45]

In both the example about conductivity and the one about fragility the situation is abnormal, and the abnormal element is an extrinsic factor: it is not part of the wire or the sheet of glass, the object whose disposition is in question. But the *bona fide* exercise of a disposition often depends on extrinsic factors, such as the air injected with the petrol into a cylinder so that a spark will cause the petrol to ignite. So (C) cannot be fixed by excluding extrinsic factors wholesale. If we add a *ceteris paribus* clause or one requiring the process from *s* to *r* to be 'non-deviant' or 'normal', the result is plausible, since we have the leeway to decide whether other things *were* equal, or whether the process from the stimulus to the response *was* normal.

(C*) A substance *x* is disposed at time *t* to give response *r* to stimulus *s* iff *x* would give response *r* if *x* were to undergo stimulus *s* at time *t*, *other things being equal*

(C**) A substance *x* is disposed at time *t* to give response *r* to stimulus *s* iff *x* would give response *r via a normal process* if *x* were to undergo stimulus *s* at time *t*

But a reductive definition of a disposition will need to be specific: it will need to specify *which* other things have to be equal, and *what* kind of process is normal. And while we can say in general terms that other things *relevant to x's disposition to give response r to stimulus s* have to be equal, and the process has to be *the kind of process that occurs when x manifests the disposition to give response r to stimulus s*, these will differ in each case. Thus, other things are *not* equal so far as conductivity is concerned when a wire is attached to the kind of safety device described above, and the process from *s* to *r is* deviant where fragility

[44] This example is adapted from one described in Bird, *Nature's Metaphysics*, 29.

[45] See Molnar, *Powers*, 91.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

is concerned when dropping a sheet of toughened glass attached to explosive device detonates the device.

So far, this is not a fatal blow to the reductionist programme. Powers could still be 'congealed hypothetical facts or states of affairs', even if no general formula covered every power. But if we cannot specify precisely when other things are equal and what counts as a normal process *for each particular case* – fragility, conductivity, the desire to please one's mother, etc. – then we are in effect still saying that so far as disposition *D* is concerned, other things are equal, or the process is normal, *if, and only if, D is manifested*. And this *is* tantamount to an admission that the reductionist programme is stymied, because (C*) and (C**) cannot be serviceable templates for reductive definitions if the left-hand side explains the right-hand side, instead of the other way around.

The debate about deviant causal chains between desires and acts, and the so far less protracted debate about physical dispositions, suggest that this how things are turning out. Several clever philosophers have tried to show how to eliminate the deviant causal chains between triggers and manifestations, but it looks increasingly as if the concepts of substance, cause and event are inadequate to account for the exercise of a disposition. Be that as it may, while the debate about dispositions is contentious, it is acknowledged on all sides that in every case where the manifestation of a disposition involves a process running from a stimulus to a response, the stimulus can cause the response *without the disposition being manifested*, since the relation between a disposition and its manifestation is process-specific whereas the causal relation between events is process-unspecific.

This is all that concerns us for present purposes, because it shows that the point on which Anscombe and Davidson are agreed, namely, that philosophers who insist that desires are causal factors 'can do no more than postulate a "right" causal connexion' between a desire and an act, is simply a corollary of this larger claim about disposition and powers in general. Hence, it has nothing specifically to do with normativity or teleology, and it has no tendency to show that when a desire explains an act, it does not do so as a causal factor. So there is no need to show how to eliminate the deviant causal chains between desires and acts in order to defend the claim that explanations of intentional action are causal explanations, as long as we do not insist that the exercise of a disposition is reducible to a specific kind of concatenation of events.[46]

[46] John Bishop claims that we need to show how to eliminate deviant causal chains in order to defend the idea that intentional acts 'consist in

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

6.

I conclude that Davidson's view that desires are causal factors is not cast into doubt by the impossibility of eliminating the 'deviant' causal chains, but the price of showing this is a much more radical departure from the Humean theory of causation than Davidson was prepared to contemplate. But if we leave the topic here, we shall miss the significance – both for the topic itself and for the twentieth-century debate – of the fact that desires are manifested in goal-directed behaviour.

As we have seen, Davidson argues that desires are causal factors by pointing out that while it is true that an explanation such as, James went to church because he wanted to please his mother, attributes a justification to the agent, it must do more than this, because it is possible to see a certain kind of value in doing something without doing it for that reason.[47] How, Davidson asks, can we turn the italicized 'and' into a 'because' in 'James went to church *and* wanted to please his mother (and knew or believed that going to church would please her)'?

The answer he favours is that 'James went to church because he wanted to please his mother' explains why James went to church causally. But he does not claim to have *proved* this. His argument is in effect a challenge:

> One way we can explain an event is by placing it in the context of its cause; cause and effect form the sort of pattern that explains the effect, in a sense of 'explain' that we understand as well as any. If reason and action illustrate a different pattern of explanation, that pattern must be identified.

---

behaviour that is caused by appropriate mental states' (*Natural Agency* (Cambridge: CUP, 1989), 2 & 148). By contrast, Erasmus Mayr claims only that we would need to show how to eliminate them in order to defend an 'event-causal' theory of intentional action (*Understanding Human Agency* (Oxford: OUP, 2011), ch.5, espec. 104. The argument here supports Mayr's view and opposes Bishop's.

[47] The basic point is anticipated by Passmore: 'explanation by reference to a "principle of action" or "a good reason" is not, by itself, explanation at all. […] For a reason may be a "good reason" – in the sense of being a principle to which one *could* appeal in justification of one's action – without having in fact the slightest influence on us.' (J. Passmore, 'Law and Explanation in History', *Australian Journal of Politics and History* **4** (1958), 275.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

> Hampshire rejects Aristotle's [conception] of wanting as a causal factor […] But I would urge that, failing a satisfactory alternative, the best argument for a scheme like Aristotle's is that it alone promises to give an account of the 'mysterious connection' between reasons and actions.[48]

I agree that wants *are* causal factors and that (3) is a causal explanation, but we should not find it surprising that Davidson's challenge failed to convince his opponents, not only because his solution faces the problem of deviant causal chains, but also because there are other patterns of explanation that apply to intentional action apart from cause-and-effect, in particular, act-and-aim.

An agent's *aim* in doing an intentional act is by definition the content of the desire manifested by the act, in other words, the content of the desire (non-deviantly-) because of which she does it. For example, Judas's aim in kissing Jesus was to betray him if, and only if, he kissed him because of his desire to betray him. (Remember: desire includes the whole gamut of wanting or valuing, as well as what we would normally call 'desire'.) If one begins an act because one wants to φ and continues it because one wants to ψ then one's aim in doing it changes; and if one does an act both because one wants to φ and because one wants to ψ, or partly because one wants to φ and partly because one wants to ψ, one's aim is compounded or divided in the same way.

As for *intentions*, philosophers tend to conceive of them as attitudes, rather than contents of attitudes, but both conceptions are equally legitimate. Like other names of speech acts or mental states, such as 'desire', 'belief' and 'assertion', we use 'intention' both to refer to a case of someone's intending something and to what someone intends. For example, 'Judas's intention in kissing Jesus was to betray him', like 'Tom's belief is that the minister will resign', refers to a content; whereas 'Judas stated (revealed, abandoned) his intention to betray Jesus', like 'Tom stated (…) his belief that the minister will resign', refers to an attitude.

Now, when we refer to someone's *intention in doing something* – e.g. Judas's intention in kissing Jesus or James's intention in going to church – we are referring to a content, again, the content of the desire manifested by the act. There is no distinction between an agent's aim in doing something and his intention in doing it. Hence,

> (1) James went to church (non-deviantly-) because he wanted to please his mother

[48] 'Actions, Reasons, and Causes', 10*f*.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

and

> (2) James's aim (intention) in going to church was to please his mother.

contain precisely the same information, so if wants are causal factors, (2) contains non-redundant causal information about James's going to church, and there is a perfectly good sense in which it is a causal explanation. However, the content of an attitude is an intensional object, an object of thought, and therefore *not* a causal factor; and the pattern that ostensibly explains James's churchgoing in (2) is not cause-and-effect but act-and-aim. So a philosopher like Anscombe, who doubts whether wants are causal factors, will not find it difficult to explain the difference between 'and' and 'because' without conceding that they are. She will simply point out that if James's *aim* or *intention* in going to church was to please his mother, it follows that he went to church because he wanted to please her – *because*, not merely *and*.[49]

So the act-and-aim scheme is quite capable of explaining the 'mysterious connection' between reasons and actions, and Davidson's challenge can be met. Anscombe's dismissive comment about Davidson quoted above – 'the solution lacks acumen', etc. – seems to pursue this line of thought:

> True, not only must I have a reason, it must also "operate as my reason": that is, what I do must be done *in pursuit* of that end and *on grounds* of the belief.[50]

(Note that '*in pursuit* of that end' and '*on grounds* of the belief' are equivalent: James's aim in going to church was to please his mother if, and only if, he went to church *on the grounds* that doing so would please her.) Her point seems to be that '*in pursuit* of that end and *on grounds* of the belief' explains the difference between the *and* case and the *because* case perfectly well without mentioning causation.

[49] See Scott Sehon, 'Teleology and the Nature of Mental States', *American Philosophical Quarterly* **31** (1994), 64; George Wilson, 'Reasons as Causes *for* Action', in G. Holmström-Hintikka & R. Tuomela, *Contemporary Action Theory I: Individual Action* (Dordrecht: Kluwer 1997), 68.

[50] Anscombe's solution is defended in Rowland Stout, *Things That Happen Because They Should* (Oxford: OUP, 1996); Scott Sehon, *Teleological Realism* (Cambridge MA: MIT Press, 2005); P.M.S. Hacker, *Human Nature: The Categorial Framework* (Oxford: Blackwell, 2007).

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

That is why I said we should not find it surprising that Davidson's challenge failed to convince his opponents. But Anscombe's exclusive attachment to the act-and-aim scheme does not seem any more astute than Davidson's to the cause-and-effect scheme. In effect, they agree that (1) and (2) contain the same information, but he prefers to explain (2) in terms of (1), and therefore claims that both explanations rely on the cause-and-effect scheme applied in (1), whereas she prefers to explain (1) in terms of (2), and therefore believes that both explanations rely on the act-and-aim scheme employed in (2).[51] But if a desire is a disposition to pursue an aim, and an aim is the content of a desire, then both explanations apply both schemes.

On the one had, (1) explains an act by identifying the desire because of which the agent did it. So if a desire is a disposition, and a disposition is a causal factor, and an explanation that identifies a causal factor applies the cause-and-effect scheme, then (1) applies the cause-and-effect scheme. But a disposition is defined by its manifestation, and a desire is manifested in the pursuit of an aim. So unlike explanations that mention simple physical dispositions such as solubility and fragility, (1) implicitly applies the act-and-aim scheme as well.

On the other hand, (2) evidently applies the act-and-aim scheme, because it explains an act by identifying the agent's aim in doing it. But if the agent's aim in doing an act is the content of the desire because of which he does it, then (2) implicitly applies the cause-and-effect scheme as well.

In sum, an explanation of an intentional act that refers to the desire the act expressed, or to the intention with which it was done, is both causal and teleological. It is causal because it refers to a disposition, and it is teleological because the kind of disposition it refers to is a disposition to pursue an aim, in other words, a disposition that is manifested in goal-directed behaviour.

Let me add a final comment about the act-and-aim scheme and the cause-and-effect scheme. Carl Ginet defends a way of explaining the 'mysterious connection' between reasons and actions without conceding that wants are causal factors, in which the explanatory scheme they are supposed to illustrate involves intentions considered as attitudes, rather than as aims.[52] He argues that the 'and' can be

[51] See Anscombe, *Intention*, 18; Davidson, 'Actions, Reasons, and Causes', 7*f*.

[52] See Carl Ginet, *On Action* (Cambridge: CUP, 1990), ch.6, and Carl Ginet, 'Reasons Explanations of Action: Causalist versus Noncausalist

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

turned into a 'because' in 'James went to church *and* he wanted to please his mother' by adding the following condition, which mentions a *de re* intention concurrent with the action it explains:

> (I) When James went to church, he intended the act to please his mother.

Notice that the intention Ginet postulates must be *de re*, because if James went to church *because* he wanted to please his mother, then pleasing his mother was not merely something he was conscious of aiming or intending to do when he set off for church, like getting home in time to watch the football on TV. It was the aim *at which his act was directed*, or *what he intended it to achieve*.

Ginet denies that (I) mentions a 'causal condition'. He says (i) that it 'does not entail that the accompanying intention it mentions caused the action'; and then, amplifying this, (ii) that it does not 'entail anything at all about what, if anything, caused the action'.[53] But whereas (i) is clearly true, since accompanying does not imply causing, (ii) is question-begging. For if Ginet is right in thinking that (I) implies that James went to church *because* he wanted to please his mother, then whether it entails something about what caused the action depends on whether desires are causal factors. In other words, the explanation-scheme (I) applies explicitly is not the cause-and-effect scheme. It is a variant of the act-and-aim scheme, which differs from it in referring to an intention-attitude instead of an intention-aim. But Ginet is not entitled to assume that (I) does not implicitly apply the cause-and-effect scheme as well.

So, instead of concluding that we *can* explain an intentional act without implying anything about its causal history, Ginet should have concluded that Davidson's challenge does not prove we *cannot*. Like Davidson and Anscombe, he seems to assume that *either* explanations of intentional action apply the cause-and-effect scheme (as his 'causalists' maintain) *or* they apply the act-and-aim scheme (as his 'noncausalists' maintain), but not both.

Accounts', in Robert Kane (ed.), *The Oxford Handbook of Free Will* (Oxford: OUP, 2002). A similar view is defended in George Wilson, *The Intentionality of Human Action* (Stanford, CA: Stanford University Press, 1996).

[53] Ginet, 'Reasons Explanations of Action: Causalist versus Noncausalist Accounts', 166.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press

7.

The argument in this article can be summarized as follows:

Intentional action does not fit the theories of causation and causal explanation that held sway in the mid-twentieth century. Desires are not causes as the Humean theory conceives of them; our knowledge about motives and reasons – especially our own motives and reasons – is not causal knowledge, as the Humean theory conceives of it; and explanations of intentional acts do not subsume them under laws. Wittgenstein and his followers were right to point these things out.

Davidson's minimal adjustment to the Humean theory of causation and Hempel's theory of causal explanation proved to be too minimal when it was faced with the problem of deviant causal chains. And his defence of the Aristotelian doctrine that desires are efficient causes was ineffective – I mean the argument about turning 'and' into 'because' – because it failed to take another Aristotelian doctrine into account, namely, that final causes play as indispensable a role in our explanations of intentional action as efficient causes.

Nevertheless, Davidson's principal claims about the explanation of intentional action were right: desires are causal factors and rationalizations are causal explanations. But as Anscombe put it in *Intention*: 'this sort of causality […] is so far from accommodating itself to Hume's explanations that people who believe that Hume pretty well dealt with the topic of causality would entirely leave it out of their calculations.'[54]

Finally, we do not have to choose between the view that explanations of intentional action refer to final causes and the view that they refer to efficient causes. For if desires are dispositions that are manifested in goal-directed behaviour, explanations that identify the desires our acts are intended to satisfy will inevitably do both.[55]

JOHN HYMAN (john.hyman@queens.ox.ac.uk) *is Professor of Aesthetics at Oxford University, and a Fellow of The Queen's College. He is joint editor (with Hans-Johann Glock) of the* Blackwell Companion to Wittgenstein, *which will be published by Wiley in 2014. A new book,* Action, Knowledge, and Will, *will be published by OUP in 2014.*

[54] *Intention*, 16.

[55] I am grateful to a number of friends and colleagues who were kind enough to comment on a draft of this article, especially Maria Alvarez, Alexander Bird, Jennifer Hornsby, Anthony Kenny, Erasmus Mayr and Kieran Setiya.

https://doi.org/10.1017/S0031819113000685 Published online by Cambridge University Press