



Title	Mechanisms that respond to reasons : an Aristotelian approach to agency
Authors(s)	Stout, Rowland
Publication date	2012-12
Publication information	Stout, Rowland. "Mechanisms That Respond to Reasons : An Aristotelian Approach to Agency." Notre Dame University Press, 2012.
Publisher	Notre Dame University Press
Item record/more information	http://hdl.handle.net/10197/5152

Downloaded 2024-05-25 12:17:17

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

Mechanisms that respond to reasons: an Aristotelian approach to agency

Section 1 - Introduction

Are there any mechanisms in the natural world that respond to reasons – that are sensitive to considerations about what they *should* do? I think that the answer is that there are approximately 6.6 billion of them on this planet alone. This is not to say that there is nothing more to being a person than being a rational agent – a reasons-responder. My claim is just that to the extent that we are *agents* we are mechanisms that respond to reasons.

The idea of being a mechanism that is responsive to reasons is thoroughly Kantian and represents us both as belonging to the phenomenal world and as being in touch with the ‘supersensible’ space of reasons. It is on the one hand an idea completely characteristic of the Age of Enlightenment – a way of defining the human condition that makes no essential reference to God. As such it belongs to a structure of ideas that Gerry Hanratty criticised for being too limited.¹ But it is also precisely this idea of being caught between two spaces – the straightforwardly physical and the transcendent – that interested Hanratty in his discussions of Gnosticism.² The central idea there was that physical material entities like us can gain knowledge of a transcendent reality and thereby regain some element of divinity that our material nature has otherwise blocked from us.

¹ Hanratty (1995).

² Hanratty (1997).

In Hanratty's discussions both Hegel and Heidegger were shown to have an element of Gnosticism in their thinking. But it would be harder to claim that Kant's conception of the will as the sort of causality that belongs to living things in so far as they are rational – as the power to act according to the idea of rules – has any such element.³ According to Kant we bring the transcendent space of reasons into a physical manifestation every time we move our bodies, and not just at moments of religious insight. Practical rationality seems to be an altogether more mundane concept than that of Sophia – that special and indeed secret insight that the Gnostics claimed for themselves.⁴

So I defend the claim that people are mechanisms that respond to reasons in this tribute to Gerry Hanratty because it is just the sort of claim that he would have tried to undermine in his characteristically unassertive way. The claim that, as agents, we are mechanisms that respond to reasons has a rather dispiriting quality to it. In particular it may strike one as both too optimistic and too pessimistic about human nature. It will seem too optimistic to the extent that the requirement that one can respond to reasons seems very demanding. Are we not also irrational creatures – still very recognisably doing things in the world when behaving in a way that fails to respond to reasons? My reply to this will be to insist on very undemanding conceptions of rationality and reasons when defending the claim that, as agents, we are sensitive to such things. Even when acting irrationally we are responsive to reasons.

³ See the *Groundwork*, 36 and 63 and the *Critique of Practical Reason*, 32.

⁴ In fact I think that practical rationality is far from the trivial calculation of preference maximisation that the proponents of 'rational choice theory' have taken it to be. I take it to be dynamic, open-ended and uncodifiable. But I do not take it to be mystical in any sense.

It will seem too pessimistic to the extent that the idea that we are mechanisms is somehow demeaning. It makes our agency seem mechanical and it makes the notion of free will problematic. If our agency is characterised by the law-governed regularity of the behaviour of a mechanism or machine then where is spontaneity; where indeed is choice? My reply to this is to point out that a mechanism that responds to reasons would be a very unmechanical sort of mechanism. The law that governs its behaviour would be the open-ended law: “Do what rationality dictates.” Assuming rationality itself cannot be codified then nor can the behaviour of such a mechanism. Part of what it is to be an agent with free will is to be capable of determining what one should do according to the reasons one has. This makes free will coincide with Kantian autonomy, and while many people still find this identification problematic I will leave it unchallenged here.

There are plenty of misleading images that we can pull down from popular imagination of what a mechanism that responds to reasons might look like. One of the most familiar is Mr Spock from Star Trek. He is an alien Vulcan logic machine incapable of emotion – very good at 3-D chess but not so good at relating to the human predicament. The image is misleading because our emotionality should be taken to be part of our rationality (and *vice-versa*).⁵ The reasons that we respond to may be ‘agent-centred’ – they may respect our very human and very particular perspective in the world. Logic – usually taken to be an impersonal structure of truth-preserving inference – need not be taken to capture rationality.

⁵ The Vulcans turn out to be a proud and loyal people – very far from being unemotional. See Solomon (1976) for a spirited – more or less Aristotelian - defence of the inextricable entangling of reason and passion.

This is all by way of preamble. I am not going to be able to argue for very much of that here. What I will be concerned with is the *causal* aspect of our agency. I think that getting this right is the key that unlocks much of the so-called problem of agency. If we can characterise our causal relationship with reasons in a way that does not leave it utterly mysterious how abstract ('supersensible') rationality can be manifested in physical behaviour then we have made some real progress. We will have shown how something that cannot be determined by empirical science (rationality) can yet be bound up in physical processes without our having to make any gestures at God.

Section 2 – A causal approach to agency

I will argue for a causal approach to agency based on an Aristotelian model of causal processes rather than on the more familiar model of a network of causal links.⁶ So I will defend the Enlightenment conception of humans causally bound up with reasons by rejecting the Enlightenment (or at least the Humean) conception of causation as a law-governed link between events in favour of the very model of causation that these Enlightenment thinkers rejected. Processes in this Aristotelian model are the realisations of potentialities. And mechanisms are the things that have potentialities; thus processes can also be thought of as the workings of mechanisms. So this approach takes the

⁶ I have been arguing for this in various ways, some of which recur here, for some time now. See Stout (1996; 2002; 2005; 2006).

working of a certain kind of mechanism to be essential to agency. That special kind of mechanism is what we refer to as an agent.

I will suggest that the potentiality that is characteristic of agency is the potentiality to respond to reasons. So my view is that agents are reasons-responsive mechanisms, to use a phrase coined by John Fischer and Mark Ravizza in their 1998 book, *Responsibility and Control*. We are agents; so that's what *we* are: reasons-responsive mechanisms.

Fischer and Ravizza claim that what is required for morally responsible agency is *not* a condition they call 'regulative control', which is the power to act differently. Rather it is what they call 'guidance control', which an agent exhibits if their behaviour issues from their own reasons-responsive mechanism (1998, 54). What they are doing here is presupposing a notion of acting for a reason and trying to work out what is required for the extra condition of acting with moral responsibility. My goal is different; it is to work out what is required for acting for a reason. So the two accounts are accounts of different things. Also there are some technical differences between the ways the accounts of reasons-responsiveness are worked out. So, although I want to mark the clear similarities between our approaches I am not setting out to defend Fischer and Ravizza across the board.

My first task is to motivate having a causal approach to agency at all. Although many people trying to construct theories of action will straightaway assume that they are aiming for a causal theory this is by no means trivial. Indeed I am going to reject the

standard causal theory of action, which identifies action with mental states or events causing physical behaviour. I think the standard causal theory gets the cause wrong, gets the effect wrong, and fails to characterize *how* the effect is brought about. So my defence of causalism will not go through this route. I want to start a bit further back with the claim that agency is best understood in terms of something - the agent - making in a certain way events happen and states obtain. In my version of a causal approach to agency the cause is the agent, not one of their mental states; the effects are worldly states obtaining and events occurring, rather than just the bodily behaviour of the agent; and what really characterises agency is how the agent brings about these things – the process.

The claim that agency essentially involves a certain kind of making things happen and states obtain is certainly not accepted by everyone, nor even by everyone who accepts some sort of causal approach to agency – see Thalberg (1967). I want to consider a two stage argument. Firstly, agency should be understood in terms of a certain kind of doing. Then, doing things should be understood in terms of making the things done happen or obtain. For example, raising one's arm is making one's arm rise. So agency should be understood as a certain kind of making things happen or obtain.

Here is the argument spelt out:

1. Agency should be understood in terms of a certain kind of doing things.
2. Doing things should be understood in terms of making things happen or states obtain.

3. Therefore, agency should be understood in terms of a certain kind of making things happen or states obtain.

The phrase, “a certain kind of doing things” is very clumsy. It is really just a placeholder for whatever turns out to be essential to agency. Commonly the *intentional* quality of our doing things is taken to be essential to its counting as a manifestation of our agency. But I am cautious of introducing intentions at this stage of the story because I am not quite convinced that all manifestations of human agency are intentional. For example when one blushes, one may be blushing for a reason. Perhaps the fact that one has just done something very embarrassing gives one a very good reason to blush. Even though the behaviour is the result of the autonomous nervous system and certainly not intentional, it is at the same time a conceptual response to the way things are – not a mere reaction to some stimulus.

The second stage of the argument is the debatable one. It might be thought that action is a paradigmatically causal concept. But it would not follow that we should seek a causal *analysis* of doing, since there may be no causal notion more basic than that of *doing* by which *doing* should be understood. No one doubts that causation itself is a causal concept, but causation should not be understood in causal terms. If doing is a very basic causal concept then a causal analysis of doing might be as futile as a causal analysis of causation.

We can see this threat clearly when we ask *what* it is that is made to happen when an agent, in doing something, makes something happen. For, if there is nothing less than the action itself that the agent makes happen then no progress is made by saying that in doing something they make something happen. Indeed the claim would clearly be false. It is certainly not right to identify eating a banana with causing oneself to eat a banana. For in that case eating a banana would be causing oneself to cause oneself to cause oneself and so on *ad infinitum* to eat a banana.

Now in the very simple case of an action that is moving part of one's body there is an obvious candidate for the thing that the agent makes happen which is not the action itself. It is the event of that part of the body moving. For example, my raising my arm might be identified with my causing *the event of my arm rising*.⁷ My arm rising is not an action – it is not the same thing as what I do. To employ the Aristotelian terminology made familiar by Jennifer Hornsby's work, it is an intransitive movement, whereas moving my arm is a transitive movement (Hornsby, 1980, 2).

Even in this simple case of an action, it might be objected that I am not really causing my arm to rise; I am just raising it. The idea of causing my own arm to rise might sound rather odd – as if I have to employ some non-standard device like autohypnosis to make the arm rise. But I think this oddness may easily be attributed to the fact that in normal cases we have a simpler way to express what happened – namely that I raised my arm.

⁷ As Coope (2007) argues, if the event of my arm rising is itself a process in a substantial Aristotelian sense then it is not clear that it is distinct from the transitive event of my raising my arm after all. So to avoid this problem we must be clear that the event of my arm rising is to be understood as the realisation of a structure of changes in the position of my arm not as the process that results in these changes.

The implication that there is some non-standard route to the arm rising when I make it rise can easily be cancelled. Suppose there was something else like a force field or an electric shock that might have made my arm rise, though it did not. Then a sensible question for someone to ask is: “Did you make that happen? And the sensible reply is: “Yes I caused my arm to rise.”

So I think it is right to identify moving one’s body with causing one’s body to move. But the real problem is that on the face of it not everything one does is moving one’s body. The example of raising one’s arm is a questionable choice as a paradigm of agency. Philosophers of action very often look at these simple acts or achievements as the targets of their analysis. But it is not at all clear that what characterizes agency are acts like this rather than more complex *activities*. Many of the things we do seem to be activities rather than achievements. Eating a banana, going for a walk, thinking about a problem, chatting with a neighbour, praying are all doings that do not seem to amount to mere achievements. Most of these things involve moving one’s body in some way, but it is not at any rate obvious that they should be *identified* with moving one’s body, even in very complex ways.

Now there are some influential arguments in favour of the non-obvious claim that anything one does is a case of moving one’s body.⁸ But even if one rejected these arguments there is a more plausible fall-back position, which is that everything one does is a case of achieving *some* structure of intransitive results. These results might include one’s body moving in certain ways; but they might also include things happening outside

⁸ See Davidson (1980, essay 3).

the body. For example we might try to identify my eating a banana with my making a structure of events happen and states obtain that are characteristic of my banana eating – i.e. roughly speaking by making the banana become ingested through my mouth into my digestive system. Perhaps my going for a walk is my causing some characteristic structure of achievements associated with going for a walk. It is not obvious exactly how to describe such a characteristic structure of achievements; but it looks like a perfectly reasonable philosophical task to work it out.

One consideration in favour of thinking of actions in terms of achievements of some sort is that with these activities it is always still possible to try but fail to do them. It would be strange, but not impossible, to try but fail to eat a banana, go for a walk, chat with one's neighbour or pray. Perhaps you gag every time you attempt to swallow a bit of banana, the wind blows you over every time you take a step, the sound of traffic means that you cannot hear a word your neighbour is saying or your mind wanders every time you start to pray. This suggests that with these activities there is something that can be successfully achieved or not. And this in turn suggests that these activities involve the achievement of goals. If eating a banana is achieving a structured goal then it is also making that goal become achieved.

So I propose that a subject, S , φ -ing can be identified with S causing a structure of results characteristic of their φ -ing . Putting this together with the first stage of the argument that what characterises agency is a certain kind of doing things, we get the conclusion

that what characterises an agent's agency is their causing in a certain way a structure of results characteristic of some activity.

Section 3 - Two kinds of causal theory

The standard way to analyse a causal notion – F – uses the following sort of schema:

S F -s O if and only if some aspect of S causes O to have some quality characteristic of F .

As a simple example, consider the wind drying the washing. Using the schema above this is to be understood as some aspect of the wind causing the washing to be dry. Which aspect of the wind should be specified here is up for discussion. Perhaps it is the wind itself that is doing the causing or perhaps it is some event or state associated with the wind. This analysis explains the causal notion in terms of a causal link between two things – a characteristic cause and a characteristic effect.

As such the analysis uses a causal chain model of causation. This model takes a basic causal process to be constituted from two particular events or perhaps complexes of events and states – the cause and the effect - and a generic relation of causation linking them. The causal world can be described according to this conception by specifying which structures of events and states are linked to one another by this relation. The causal world will look like a huge chain or network of events and states linked together.

A diagram of such a causal world will have names for the events and states and a complex structure of arrows linking these names.

But this is certainly not the only model of causation. There is also an Aristotelian model that allows in addition to a cause and an effect not just a generic relation of causation linking them, but individual causal processes. In this model we can ask what the cause is and what the effect is, but in addition we can ask how the effect results from the cause – i.e. what the mechanism is. And this how-question is not to be answered just by introducing more links in a causal chain. It is answered by specifying the potentiality whose realisation is the process that the effect belongs to.

This Aristotelian model of the causal world cannot be fully pictured as a network of arrows, since each arrow would have to represent a different sort of causal process. It is a world of mechanisms having potentialities which are realised in the workings of the mechanisms and which realise each other.

The obvious inadequacy of the causal chain model as an account of the wind drying the washing is that by only specifying the cause and the effect it misses out the way the effect is caused, and this way too has to be characteristic of wind-drying. This can easily be seen by constructing a deviant causal chain counterexample. Suppose the wind blew open the window, which banged into the button that turned on the tumble dryer in which the washing soon dried. The wind caused the washing to be dry but it did not dry the washing. This is because it caused the washing to be dry in the wrong sort of way. It did

not result in the washing becoming dry through the realisation of a wind-drying mechanism.

This explanation suggests an alternative way of analysing causal notions – the causal *process* model.

$S F$ -s O if and only if O having the F -quality belongs to a process that is the working of an F -mechanism (the realisation of its F - potentiality) embodied by S .

For example, the wind dries the washing if and only if the washing being dry belongs to the working of a drying mechanism embodied by the wind. “ F ” appears on both the left and the right hand side of this biconditional, which means that this is not a reductive account. But there is no vicious circularity here. The characteristic results that go with a certain kind of causal notion can be established at least partially independently, and then we can establish on any particular occasion whether something belongs to a process that is the working of a mechanism with these characteristic results. For example, the characteristic results of the mechanism of the wind drying washing are something like this: the longer the washing is in the wind and the stronger the wind the dryer it becomes, until it is completely dry. If the drying of the washing belongs to a process that is the working of a mechanism with these characteristic results embodied by the wind, then we can say that the washing was dried by the wind.

In the deviant causal chain example the washing becoming dry does not belong to the working of such a mechanism. There is a drying mechanism in this example, but it is not embodied by the wind. And the way the relevant bit of the mechanism that is embodied by the wind works does not have these characteristic results. In general the causal process model is immune from deviant causal chain counterexamples precisely because it does not employ the idea of a causal chain.⁹

However this account does employ ideas that many modern philosophers of causation find deeply suspicious. First of all is the idea of a mechanism. A mechanism is identified by its potentiality. Here I am using the Aristotelian notion of a potentiality. Aristotle said: “Motion is the fulfilment of what exists potentially, in so far as it exists potentially” (*Physics*, 201a10-11). So a process is the realization of a capacity or disposition for certain results in certain circumstances. You have to characterize a structure of stages to specify the potentiality. But what is required for the process to be happening is not just that that structure of stages is in train, but that there is a potentiality for such a structure and that this potentiality is being realised.

There are two sets of conditions associated with something having a potentiality. There is the set of statements that describe what the potentiality is a potentiality for. And there

⁹ Some of the classic examples which get included in the deviant causal chain discussions do not strictly speaking employ deviant causal *chains*. For example, Chisholm’s original (1966) example of Joe running down his uncle by mistake in his hurry to drive to the spot where he could kill his uncle can be adapted to one where Joe kills the pedestrian on purpose but still does not intentionally kill his uncle. Here the process resulting in Joe’s behaviour is properly characteristic of agency; it is just that it is wrong to describe this behaviour as intentionally killing his uncle.

is that underlying nature or set of conditions that grounds the potentiality.¹⁰ These latter conditions are divided into two categories – the conditions that constitute the potentiality or mechanism itself and the conditions that constitute the realisation of the potentiality or the operation of the mechanism. Usually this distinction is made in a pragmatic way, and for simplicity we can lump them all together as the underlying conditions of the process.

One worry people have with this sort of account is that introducing potentialities and mechanisms provides *empty* explanations. Citing the dormitive virtue of opium has little explanatory value when trying to explain why opium puts you to sleep. Likewise, saying that the wind dries the washing if and only if the washing becoming dry belongs to a process which is the realisation of a drying mechanism embodied by the wind does not have much explanatory power.

Now I do want this causal process account of causal notions to apply to the claim that opium put a patient to sleep. I want to say that it is true that the opium put the patient to sleep if and only if the patient falling asleep belonged to a process that is the realisation of a dormitive virtue or potentiality embodied by the opium. I also want to say that this claim is explanatory. But it does not explain why or how the opium put the patient to sleep; it explains what it *means* to say that the opium put the patient to sleep. So no wonder it has the sound of an empty tautology. Also it is not the end of the explanation but only the start. For one must go on to explain what a dormitive potentiality is – what the characteristic results of such a process are. And finally, although neither the claim

¹⁰ Philosophers opposed to this talk of mechanisms and potentialities are also suspicious of this idea of a grounding relation, and certainly the Aristotelian approach owes a treatment of it, even though I will not attempt one here.

that the opium put the patient to sleep nor what the causal process model takes to be the equivalent claim about the realisation of the dormitive potentiality explain how the opium put the patient to sleep, they both do explain why the patient fell asleep.

Section 4 – A causal process account of agency

Let us go back to agency. In section 2 I argued that we should start by trying to understand the special way an event is made to happen or a state is made to obtain in action. More precisely, we should start by understanding what it is for something to cause things to be *G* in the way that is characteristic of agency, where *G* stands for the intransitive characterisation of what they are achieving. A typical way for the causal chain model to explain this is the standard causal theory of action, which works in the following sort of way, though of course there are many variations of this:

An agent causes things to be *G* in a way that is characteristic of agency if and only if the agent's intending to cause things to be *G* causes things to be *G*.

This is vulnerable to the deviant causal chain objection I raised in section 3. The Aristotelian causal process model looks instead for a way to describe the sort of mechanism whose working constitutes the process of acting. Harry Frankfurt has been an important proponent of this sort of approach.

[T]he state of affairs while the movements [of a person's body] are occurring is far more pertinent [than the causes from which they originated]. What is not merely pertinent but decisive, indeed, is to consider whether or not the movements as they occur are under the person's guidance. It is this that determines whether he is performing an action. Moreover, the question of whether or not movements occur under a person's guidance is not a matter of their antecedents. Events are caused to occur by preceding states of affairs, but an event cannot be guided through the course of its occurrence at a temporal distance. (Frankfurt 1978, 158)

For Frankfurt, what distinguishes guided action is the causal mechanism not the causal antecedents. The way this mechanism works must be characteristic of an agent causing things to be *G*. What the characteristic results are such a mechanism is a central question of the philosophy of agency. In 1964 Charles Taylor gave the following account of the teleological explanation of action:

To offer a teleological explanation of some event or class of events, e.g., the behaviour of some being, is, then, to account for it by laws in terms of which an event's occurring is held to be dependent on that event's being required for some end. To say that the behaviour of a given system should be explained in terms of purpose, then, is, in part, to make an assertion about the form of laws, or the type of laws which hold of the system. (1964, 9)

I think he is right; a mechanism whose working is characteristic of an agent achieving things is one that involves sensitivity to reasons – in particular teleological reasons.

When we do things – cause things to be a certain way – we do them for reasons. And very often these reasons are teleological; we do things for the sake of other things.¹¹ For example, I walk to the letterbox because that is what I should do in order to post the letter.

So I am embracing a powerful Kantian tradition in the philosophy of action in which what is special about the way things are made to happen in action is that they are made to happen in a way that is sensitive to reason. For this to be even plausible, a couple of qualifications are required. Firstly it is possible to act but not for any particular reason. I might eat a banana but not do it because I was hungry or even because I particularly felt like it. I just chose to eat a banana and went on to eat it. But even if the action as a whole has no reason, the process of acting is still sensitive to reason. The top-level goal of eating a banana may just be given; but my behaviour is still sensitive to what I should do to achieve it.

Secondly it must be possible for quite irrational behaviour still to count as done for a reason. Somehow we must allow for behaviour that is rational within an irrational system. This behaviour may be irrational absolutely, but relative to that system it is

¹¹ It is precisely this requirement that the reason for action must have a teleological structure that should be loosened if we are to broaden this account to include non-intentional emotional responses like blushing considered briefly earlier.

rational. My suggestion for accommodating this is to require that the behaviour that accords with reason does so relative to a *version* of practical rationality. A version of practical rationality is to be construed as a way of making recommendations for action. Perhaps it must involve some sort of means-end rationality, though this might be questioned. But it may also embed certain goals which need not be justified by this or any other version of practical rationality, as well as certain assumptions which need not be justified either and indeed may simply be false. For example the goal of eating a banana may be fixed in a version of practical rationality, and that version might also embed the quite false assumption that there is a banana in the kitchen. According to this version of practical rationality, and assuming that there is no better way to get a banana and so on, going to the kitchen is the thing to do. This goal of going to the kitchen is not fixed in that version of practical rationality but derived from applying it to the agent's environment. In general, a version of practical rationality, although it may embed certain fixed assumptions and goals must also involve some sort of sensitivity to the way things are.¹²

So when I talk about a mechanism resulting in what 'should happen' I am using a relative notion of 'should'. It results in what should happen according to some version of practical rationality. The reasons the mechanism responds to are reasons according to some version of practical rationality. We also use the words 'should' and 'reason' in an absolute way. We might say that really there is no (absolute) reason to go to the kitchen, since there is no banana there. But we are not responsive to such absolute reasons.

¹² See Stout (1996 and 2006) for attempts to work this out in some detail.

Fischer and Ravizza (1998), in their account of reasons-responsive mechanisms appear to be working with an absolute conception of reasons. This leads to far too demanding a requirement of agency where to be a morally responsible agent your behaviour has to fit the contours of absolutely correct reasoning perfectly. There is no scope here for more or less irrational morally responsible agents. Fischer and Ravizza, of course realise that this will not do, and they concede that the responsible agent's behaviour only has to be loosely or 'moderately' responsive to absolute reasons. But this makes their account very messy with *ad hoc* complications introduced to deal with a series of reasonable counter-examples.

Instead of requiring only moderate sensitivity to absolutely correct reasons Fischer and Ravizza might have required actual sensitivity to only relatively correct reasons. We do not require every agent to be in touch with some perfect version of rationality. Agents can get things wrong even when they are acting for reasons. The reasons for which you act need not be reasons that would figure in a perfect version of rationality, but just reasons that figure in some – possibly quite flawed – version of practical rationality. It is not obviously false that non-human animals do things for reasons. But if other animals do things for reasons, what count as reasons for them will do so only according to a very limited version of practical rationality.

So my proposal is that the special kind of mechanism that is characteristic of agency is simply a mechanism that characteristically results in what should happen according to

some particular, and possibly flawed, version of practical rationality. This gives us the following causal process theory of action:

Suppose that the result characteristic of something, S , φ -ing is G . Then, S φ -ing counts as an action if and only if:

- (i) there is a version of practical rationality that recommends in the circumstances that G is to be achieved;
- (ii) G being achieved belongs to a process that is the working of a mechanism embodied by S which characteristically results in what should happen according to that version of practical rationality.

Section 5 – We are the mechanisms

If I am right then agency is characterised by the working of a certain sort of mechanism – a mechanism that works by making happen what should happen (according to some version of practical rationality). But we, all 6.6 billion of us, are agents. That does not mean that we contain agency inside us. It means that we are identified *as* agents; what we are are things that act. So the special mechanism that responds to reasons and that characterises our agency, on this Aristotelian picture of action, is not just part of us; it is us. We are such mechanisms. This is not to deny some essential place for consciousness in our conception of ourselves. For a mechanism that is sensitive to reasons must be sensitive to the world. You cannot be the sort of mechanism that makes happen what should happen unless you are *aware* of the world around you. In virtue of being agents

we are also subjects.

References

- Chisholm, R. 1966, "Freedom and Action", Lehrer, K. (ed.), *Freedom and Determinism*, New York: Random House, 28-44.
- Coope, U. 2007, "Aristotle on Action", *Proceedings of the Aristotelian Society* Supplementary volume LXXXI, 109-37.
- Fischer, J. and Ravizza, M, 1998, *Responsibility and Control*, Cambridge: Cambridge University Press.
- Frankfurt, H. 1978, "The problem of action", *American Philosophical Quarterly* 15, 157-62.
- Hanratty, G. 1995, *Philosophers of the Enlightenment*, Dublin, Four Courts Press.
- Hanratty, G. 1997 *Studies in Gnosticism and in the Philosophy of Religion*, Dublin, Four Courts Press.
- Hornsby, J. 1980, *Actions*, London: Routledge.
- Solomon, R. 1976, *The Passions*, Garden City NY: Anchor Press/Doubleday.
- Stout, R, 1996, *Things That Happen Because They Should*, Oxford: Oxford University Press.
- Stout, R. 2002, "The right structure for a causal theory of action", *Facta Philosophica* 4:1, 11-24.
- Stout, R. 2004, "Internalising practical reasons", *Proceedings of the Aristotelian Society* 44, 229-43.
- Stout, R. 2005, *Action*, Teddington: Acumen.
- Stout, R. 2006, *The Inner Life of a Rational Agent*, Edinburgh: Edinburgh University Press.

Taylor, C. 1964, *The Explanation of Behaviour*, London: Routledge.

Thalberg, I. 1967, "Do we cause our own actions?", *Analysis* 27, 196-201.