

I, myself, move

Lucy O'Brien 

Department of Philosophy, University College
London, London, UK

Correspondence

Lucy O'Brien, Department of Philosophy,
University College London, Gower Street,
London WC1E 6BT, UK.
Email: l.obrien@ucl.ac.uk

Abstract

This paper addresses the question “what connection is there between our answer to the question of what we are, and the question, what our actions are?” Suppose that actions are reflexive changes of agents. On that supposition, there would be a direct connection between the answers to those two questions. An action of mine will be a reflexive change of me, and what I am will fix the nature of those changes. I hold that supposition to be true and consider reasons in favor of believing it. However, the paper is not primarily aimed at defense of that thesis. It rather concerned with exploring what consequences accepting it has for the competing notions of what we are, given what we ordinarily think actions are, and bringing to light a tension between thinking of actions as reflexive changes of agents in this way, and a kind of causal understanding of actions that is prevalent. What emerges is that we should shift where we start our theorizing: we cannot assume that action theory primarily involves the task of characterizing the relation between an agent and changes caused, rather than a characterization of a particular kind of relation between the agent and herself.

[...] I manifest self-conscious thought [...] not in knowing which object to act upon, but in acting. (I do not move myself; I myself move.)

(Evans, 1982, p. 207)

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Authors. *European Journal of Philosophy* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

I aim, in this paper, to take on very directly the question that motivates this special issue on “Beings and Doings.” The question, I take it, is “what connection is there between our answer to the question of what we are, and the question, what *our actions* are?” What if we suppose that actions, in their most basic abstract characterization, are reflexive changes of agents? By that I mean, what if we suppose that in acting an agent always changes itself, and that those changes are its actions. On that supposition, there would be a very direct connection between the answers to those two questions. If actions always involve changes to the agent by the agent, then the nature of an action will be determined by the nature of the agent. Whatever else it is, an action of mine will be a reflexive change, and what I am will determine the nature of those changes.

I hold the supposition that actions are reflexive changes of agents to be true, and will canvas some reasons to believe it. However, this chapter is not aimed at full defense of the thesis. It is primarily concerned with bringing the thesis into view, exploring what consequences accepting it has for the competing notions of what we are, given what we think actions are, and bringing to light a tension between thinking of actions as reflexive changes of agents in this way, and a kind of causal understanding of actions prevalent in much contemporary action theory.

What emerges is that the suggestion that actions are reflexive changes of agents brings with it a recommended shift in where we should start our theorizing as action theorists: we cannot assume that action theory primarily involves the task of characterizing the relation between an agent and changes that are the results of her actions, without first offering a characterization of a particular kind of relation between the agent and herself. That is not to say that many of our actions are not also *transactions* that involve changing objects that are distinct from the agent, nor that there is no account to be given of the relation between the changes to the agent, and the changes to other objects. It is rather to say that none of our actions are only transactions, and all our actions are self-changes. We change objects distinct from ourselves only by changing ourselves, and we do not understand agency unless we understand the power to reflexively change.

To a significant extent, this paper constitutes an exercise in ground preparation: if the suggested connection between the answer to the “who are we that act” question, and the answer to the “what are our actions” question is right, we should start the business of trying to understand what actions are in a way different from the way we have been. We will certainly not finish—but the task will be identified and delineated, and the ill fittingness of the structure of standard accounts will, I hope, emerge.

Having suggested that we take actions to be reflexive changes of agents I will identify an explanatory burden we face if we accept the suggestion, but do not take us, as agents, to be human animals. We may not have any philosophical theory ready to hand, but most of us think we, very often, know an action when we see one. When we see a fellow human walk into a room, sit at a table, and pick up a glass, we take ourselves to come to know that someone, walked, sat and picked up a glass. But what we see in such a case are movements of the human animal. I will argue that anyone who denies that we are human animals—who takes us to be Cartesian, Humean, or other non-human selves—but accepts that actions are reflexive changes will face the burden of explaining away such common-sense identifications of actions.

Following an examination of how the Cartesian, Humean, or brain-based Lockean, might meet that burden, we will note a striking parallel between the kind of maneuvers that suggest themselves, and recent offerings in action theory from theorists not in any way officially committed to denying that we are human animals. This raises the question of whether such accounts eschew a metaphysics of the self in their discussions of personal identity, which is nevertheless retained in the structure of their theories of action. I will argue that it is not primarily a commitment to a problematic metaphysics of the self that causes the problem. It is rather a problematic form of causal relationalism—concomitant with such a metaphysics, and perhaps motivating such a metaphysics—on which action is construed as a causal relation between a subject and her bodily movements. We should instead, I argue, take the action to be the animal movement, and take the animal movement to constitute a reflexive relation in which the animal stands in a relation to herself. This suggests that our strategy as action theorists should be a “vertical” one rather than a “horizontal” one. We should not be looking at actions and asking how an agent needs to relate to them in

order for them to count as actions, we should be looking at agents, and asking how the way they are organized and constituted allows them to relate to themselves in a way that enables them, reflexively, to change.

2 | THE REFLEXIVITY CLAIM

What must be true if there are actions? Well, one thing that seems indisputable is that there must be agents (Call this the “Agent Claim.” See O'Brien, 2017). Any action of mine, or anyone else's, could not have occurred without my, or them, existing: there are, and could not be, “unowned actions.”

But is it indisputable that there be actions without being agents? What makes it so plain that I must exist if I am to act, and that there are no “unowned actions” possible or conceived. Of course, we can contemplate the possibility that there are no actions—that they are part of the illusory face of a manifest image in which no such *psychologia* exist. Perhaps, there are only fundamental particles organized over time, more or less, action-wise. But if we think that there really are actions, then we are committed to thinking that there are agents. There is no stable view on which actions are part of our inventory, while agents are not, even if there are inventories on which there are neither. One kind of answer would be that actions are caused by agents, and that the existence of an agents is a necessary causal condition on the existence of actions. However, if we take causation to involve a relation between distinct existences then, while you would not normally get action without agent, it would not obviously be impossible. Perhaps it would be impossible that there be a world in which there were only actions and no agents, but it would not be impossible that the occurrence that is the action could occur but without the agent—it is possible that the particular occurrence that is actually the action might have been caused by something other than the agent. It might be said that were that occurrence to be caused by something other than the agent then it would not count as an action—that to be classified as an action is to be an occurrence with a given causal history involving an agent, and stripped of that history we have the occurrence, but no action. I consider the modal dependence of actions on agents as stronger than this: there is no such occurrence, action or otherwise, in a possible world where there is no agent. However, I do not intend to dispute that a necessity of causal origin thesis might have resources with which to answer the question as to why there are no unowned actions. What I do want to do is suggest that a simpler and clearer answer to that question lies in the thought that actions are always conditions of agents. If we take actions to be changes by the agent to herself, the action will be a condition of the agent, and it will be implied straightaway that the agent must exist for the action to exist: that which changes must exist for the change to it to exist.

One might worry that focusing on seeing actions as changes to the agent leaves out the primary purpose and importance of actions, and that is that they involve interactions with the world. We do not just, ourselves, change in acting; we change the objects and the environment around us. We move the kettle to the stove, the sandwich to our lips, and the note to our friend. We act to change our world—not ourselves. As said above, it is true that many of our actions are *transactions* and involve changing objects that are distinct from the agent. It may also be that it is true that the goals which we act are as often, indeed usually, concerned with changing the state of the world, rather than the state of ourselves. But we must act with ourselves; none of our actions are only transactions. We change objects distinct from ourselves only by changing ourselves. I move the kettle to the stove, the sandwich to my lips, the note to my friend, by, *myself, moving*. My power to change things distinct from me rests entirely on my power to change myself when I act. I can only change what we might call the “non-me world”—those things and stuffs distinct from me—by reflexively changing. So, even if some non-me changes are actions, they also are, or depend on—depending on how we individuate changes—self-changes.

If this is the right way to view the nature of action it makes the problem of action fall within a set of “reflexivity problems.” The difficulty of explaining the particular ways in which a subject reflexively relates to herself in consciousness, reference, knowledge, or concern, constitutes the heart of some of our most intractable, and seductive, philosophical problems. These problems are most simply characterized in relation to what we can call “reflexivity questions”:

- The problem of self-consciousness—what is the nature of a subject's reflexive consciousness?
- The problem of first person reference—what is the nature of reflexive referring by a subject?
- The problem of self-knowledge—what is the nature of a subject's reflexive knowledge?
- The problem of self-concern or self-love—what is the nature of a subject's reflexive concern or love?

My view is that the problem of action is the problem addressed by answering the reflexivity question: what is the nature of reflexive changing by a subject? These reflexivity problems have a common property—for any particular subject who is reflexively conscious, who reflexively knows, refers, cares, changes, there is no further question as to who the subject is reflexively conscious of, knows of, refers to, or cares about. The object of any one of these reflexive phenomena is their subject, because they are *reflexive*. As Tom Lehrer might put it, a reflexive attitude is “like a sewer: what you get out of it depends on what you put into it.” This, of course, means that discussion of these problems has co-existed with many competing and incompatible accounts of the nature of the subject “you put into it.” This might make it seem as though these problems are neutral with respect to different accounts of the nature of the self-conscious subject, self-referrer, self-carer/lover, self-changer, but they are not. How we think about what we are will need to be tempered by the fact that we are subjects that are capable of being reflexively conscious, of reflexive referring, of reflexive caring, and reflexive changing. And, more importantly, for what I want to explore in this piece, how we understand the phenomena in question, and the resources we have to answer the question they pose, will be significantly affected by how we understand the nature of the subject in question. Suppose, for example, that we accept the proposal made that all my actions are at least self-changes. Then, to know what my actions are, I need to know what a self-change is, and to know that—since a self-change is a reflexive changing by me—I need to know what I am. If I am an “S sort of a thing,” and if S sorts of things change in S sorts of ways, then an action will be an “S sort of a change.”

To see what the consequences of thinking about things in this way are, consider the following three claims, and the relation between them.

The reflexivity claim: Whenever an agent acts, she reflexively changes and the agent's actions are those reflexive changes.

The animal changes claim: This animal (the one sitting on this chair as I type) changes when I act, and these animal changes are my actions.

The animal identity claim: I am this animal.

I have offered some reasons in favor of considering the reflexivity claim. Accepting the claim makes sense, in a simple way, of why it is impossible for an act to exist without its agent, and is consonant with the thought that any action the subject carries out will involve some change to the subject carrying it out: we do not change the behavior of objects distinct from ourselves at distance, we also have to be involved, by changing.

The animal changes claim is a claim based on taking our ordinary practice of action perception and identification at face value. We take ourselves to know roughly what actions we, and others, carry out: we take there to be walkings, sittings, pickings up of glasses, passings of notes, and so on. Moreover, we ordinarily think of these as movements of human animals, and as movements determined by the nature and constitution of human animals. If the human animal, or the animal we associate ourselves with, had four legs, or no arms, our walkings would be quite unlike our walkings, and there may be no passings of notes.

It is obvious that the animal changes claim, and the reflexivity claim, can constitute premises in an argument for the animal identity claim, the claim that “I am an animal” (see Geddes, 2016 for related ways of arguing). If the animal movements that occur when I act are both my actions, and my actions are *reflexive* changes, then I must be the animal whose changes are those movements. If we do not think that conclusion is true we face, I think, the choice of either denying that my actions are those animal movements we usually identify as my actions, or denying that my actions are properly thought of as reflexive changes to the things that I am.

I want now to consider what strategies might be adopted by someone who denies that I am an animal but wants to hold on to the reflexivity claim. If someone holds that the agent who acts is not the animal whose movements we ordinarily take to be the actions of an agent, how should they think about the actions of that agent?

3 | CARTESIAN SELF-CHANGES AND ANIMAL CHANGE

Suppose that we assume that the thing that I am is a Cartesian non-extended mental substance. And suppose we accept the reflexivity claim, and so accept that in acting I self-change and that those self-changes are my actions. It then follows that in acting I change the Cartesian mental substance that I am.

But now we face a puzzle: The Cartesian non-extended mental substance that self-changes, is distinct from this animal. This animal changes when I act, and ordinarily we call these animal changes my actions. We go on as if the animal changes claim is true, but it cannot be if I am a non-extended Cartesian substance, and if the reflexivity claim is true. So, what are we to say of the animal changes that we ordinarily take to be my actions?

This puzzle is of course not news. There are a number of strategies of which the Cartesian can avail to try to meet it. Let me set out three types of solution:

Solution 1: Actions are mental self-changes that are *causes* of animal change: The changes to this animal are *not* my actions; they are rather only the effects of my actions. Speaking *loosely*, we call the animal changes the action, but *strictly* speaking they are not. Only the changes in the Cartesian mental substance—cogitations that are acts of will—are strictly speaking my actions.

Solution 2: Actions are either mental self-changes or are the *effects* of mental self-changes: The changes to this animal are my actions—because actions come in two kinds: those that are mental self-changes and those that are the effects of mental self-changes. The animal changes are my actions but only in virtue of being caused by self-changes that are non-derivatively my actions.

Solution 3: Actions are reflexive causings of change to the mental substance, or causings of animal change, by means of such reflexive causings. The mental substance has the power to cause: the action is an execution of that power and is not the change, but is rather the causing of the change. My actions are the causings either of reflexive changes or causings of animal change, by means of such reflexive causings. (On 'causings' See Alvarez, 2024, Alvarez & Hyman, 1998, Steward, 2012).

How exactly the solution works differs in each case, but the broad strategy is the same. There is a mental act which we can see as a reflexive change (or the bringing about of a reflexive) change and there are animal changes that stand in one kind of causal relation or another to the mental act. The animal changes either derivatively count as actions, or are not actions but are—speaking loosely—called actions.

4 | HUMEAN SELF-CHANGES AND ANIMAL CHANGE

Suppose that we assume that the thing that I am is a Humean bundle of psychological phenomena. And suppose we accept that in acting I self-change. It then follows that in acting I change the Humean bundle of psychological phenomena that I am.

But now again we face a puzzle: the Humean bundle, that self-changes, is distinct from this animal. This animal changes when I act, and ordinarily we call these animal changes my actions. So, what are we to say of the animal changes that we ordinarily take to be my actions? The Humean bundle theorist can offer slightly altered versions of the three types of solution set out above: the self-changes to the Humean bundle are my actions, and they cause animal changes; that actions come in two kinds—non-derivative self-changes and casually derivative animal changes; my actions are the causings either of reflexive changes or the causings of animal change.

In fact, things may be yet more complicated for the Humean: there is a kind of Heraclitan puzzle that faces the Humean under some accounts of what the bundle is. If, as we have assumed, in acting I self-change, then in acting I change the bundle of psychological phenomena that I am. The only way to change the bundle is to add or take away from the set of psychological phenomena that constitute it; in acting a new set of psychological phenomena come to constitute the bundle. However, if the identity and character of a bundle is given only by the psychological events that constitute it—as the identity of a set is given by the identity of its members—then there is no possibility of the bundle changing itself: it can only change into another bundle, and so another self. On this conception of the identity of the bundle, if actions are psychological self-changes, then the action, which is the changing, cannot be a constituent of the bundle. Call A the self-changing act at t. Suppose B at t is the set of psychological phenomena (P1, P2, P3...PN). If A is one of P1, P2...PN then B does not change, so A is not a self-changing act. If A is not one of P1, P2...PN, then A is not part of the bundle, so is not a self-changing act by the bundle.

5 | BRAIN SELF-CHANGES AND WHOLE ANIMAL CHANGE

The problems identified above affect not just the Humean, but most non-materialist versions of Lockeanism. The very notion of a self-change is problematic for any purely constructive or non-substantial view of the self, at least in so far as the change is taken to change the identity of the self that is supposed to be self-changing.

Setting aside the specific Heraclitan puzzle that faces the Humean, or Lockean, problems parallel to those raised for the Cartesian and Humean above also face more contemporary materialist conceptions of the self. This is notably true of what we can call “brain based” views of the self. There is a widespread tendency across contemporary philosophy, psychology, and neuroscience to talk as though the self—what I am—is identical with or constituted by my brain. So, the self-change thesis about action has the consequence that actions are changes my brain makes to itself. There are no doubt reflexive changes in the brain when I act, but saying only that leaves us with the puzzle about the more extended animal movements we have been looking at. The brain that self-changes is, at least normally, distinct from this animal that changes as a whole, and whose changes we usually call my actions. The theorist of the brain-based view of what I am can, of course, when faced with this puzzle, revert to the strategies we have considered. In particular, they can lift pretty much wholesale the strategies we offered to the Cartesian, replacing only the notion of a Cartesian self with the notion of a human brain. Thus, they can offer the following solutions:

Solution 1: Actions are brain self-changes that are *causes* of macro animal change: The macro changes to this animal are *not* my actions; they are rather only the effects of my actions. Speaking *loosely*, we call the animal changes the action, but *strictly* speaking they are not. Only the changes in the brain are strictly speaking my actions.

Solution 2: Actions are either brain self-changes or are the *effects* of brain self-changes: The macro changes to this animal are my actions—because actions come in two kinds. Those that are brain self-changes and those that are the effects of brain self-changes. The macro animal changes are my actions but only in virtue of being caused by self-changes that are non-derivatively my actions.

Solution 3: Actions are either reflexive causings of change to the brain, or causings of animal changes by means of causing changes to the brain. The brain has the power to cause changes to itself, and by such means to the whole animal. The action is an execution of that power and is not the change, but is rather the causing of the change. My actions are the causings either of reflexive changes or the causings of animal change, by reflexive changes.

6 | ANIMAL SELF-CHANGES

It should by now be obvious that we may avoid the puzzle, that arises from our taking actions to be animal changes, if we take ourselves to be identical with the animal whose animal changes we are taking to be actions. If we assume that I am a human animal, and that in acting I self-change, then in acting, I change the human animal that I am. There

is no parallel action problem here. The human animal that self-changes is identical to this animal. This animal changes when I act; we call these changes my actions. We should note three potential complexities, even on the animal view.

First, if we think of actions as reflexive *changings*, or *causings*, rather than changes, as for example Alvarez does (2024), then even on the animal view it is not settled how to think about the animal movements. The animal movements may be construed as self-changings, or self-causings, and well as changes, that are the agent's actions. Or they may be construed as changes brought about by the agent's distinct action of self-changing. The former option effortlessly delivers a fit with the naïve idea that actions are animal movements. Things are a bit less clear if we take the action not to be the movement, but only the (non-derivative) bringing about of the animal movement by the animal. It is less clear, partly, because it is not clear exactly how to think about changings and causings in relation to the changes and effects they bring about. However, even if the fit with the naïve identification is slightly more complicated on this view—it is simpler than the view on which the bringing about of the animal movement is derivative and must be understood via the bringing about of a mental change. I am going to set this complication aside and continue to talk of self-changes.

Two, a distinction is often made between the human animal and the human body. Consider the following question: are those candidate actions—the changes we have called *animal* changes—actually animal changes, or are they rather *bodily* changes? The distinction between the *human animal* and the *human body*, familiar from philosophical discussions of personal identity, is one on which the latter is usually held to have some kind of constitution relation to the former, and to change through the course of a single human animal's life. (I am not going to summarize the debate in relation to personal identity, but see Snowdon, 2014 for a recent discussion of this distinction and its implications.) This distinction then allows for the thought that while I am this human animal, I am not this human body. We then face the related task of figuring out the relation between animal changes and bodily changes, and determining what, in our story of the way they relate one to the other, we are to count as the action.

There are two ways we might accept the distinction between human animal and human body and yet try to set aside this puzzle about animal changes. First, we might accept that there is a distinction between animal change and bodily change, but deny that human actions either *are* bodily changes, or are the *causes* or *causings* of bodily changes. We might, rather, stick to the view that since I am this animal, and since actions are reflexive changes, my actions must be those animal changes. If there is a relation between those animal changes, which are my actions, and distinct bodily changes, the relation will not be one of identity or cause—but rather one of realization. The animal changes that are my actions may be realized by bodily changes, in the way that the animal I am may be constituted by this distinct body.

The second way, we might accept a distinction between animal changes and bodily changes, but try to set aside the puzzle of animal changes, is by weakening the reflexivity claim. We may allow that my actions can include changes not only that that I am identical with, but also changes to that thing which constitutes me. We may accept something like:

The reflexivity and constitution claim: whenever an agent acts, she reflexively changes, and that which she constitutes, or constitutes her, changes, in concord. The agent's actions will be those reflexive changes, and those changes, to that thing she is constituted by, or constitutes.

I do not intend, here, to get into the metaphysical complexities of how to individuate changes, or how to understand the relation between changes to that which is constituted and that which constitutes it. But we can note that when the relation proposed is constitution—as it is between the human body and the human animal—rather than containment, or causation, it is more natural to identify changes to the constituted with changes to the constituting. Identifying the shrinking of the statue with the shrinking of the material that constitutes, but is non-identical to, it seems in order in way that identifying the shrinking of the statue with shrinking of part of it, or the shrinking of the cause of its shrinking, is not. Identifying the human animal movements, that we ordinarily take to be actions, with the human bodily movements of a human body purportedly constituting the human animal, seems plausible in a way that identifying

human animal movements with movements of its brain, or with changes which causes them, does not. Of course, as the modally and historically sensitive properties of an entity, and that which constitutes it, are different, and as the human body and the human animal are distinct on this account, an explanation will have to be provided of what allows a change to A (the animal) to be the same change as a change to a distinct B (the body). If we distinguish between the human animal and human body, and are also skeptical of the prospects of a story allowing an identification of animal and bodily changes, we seem to face a choice between either (i) taking both the human animal changes, and the human body changes, to be actions—and so allow that we in some sense act twice, or (ii) identifying the action with either one of the human animal change or the human body change. If we choose option (ii) and identify the action only with the human animal change we are back to the first position canvassed above. If we choose (ii) and identify the action with the human body change we face a version of the animal changes puzzle.

Of course, if we accept, as many do, that the human body *just is* the human animal—if neither can be said to survive or change without the other surviving or changing in the same way—then there is no related animal changes puzzle that needs solving. If there is no distinction between human animal and human body then our problem disappears.

The third potential complexity arises—from another distinction. Rather than a worry about there being a distinction between human animal and human body, one might worry that making sense of self-movement requires making a distinction between the human animal and “life” or “self-conscious life.” One might agree with the theoretical claim that we need to think of actions as exercises of an agent's reflexive capacities, and agree that the walkings, sittings, and talkings, are the actions of the agent, but resist the idea that the agent that changes in so acting is the human animal, rather than the “life.” Perhaps Rödl (2007) takes this view.¹

Again, there are two ways we might respond if we recognize the distinction, but want to try to set aside our puzzle. First, we might accept that there is a distinction between animal changes and “life” changes, but deny that human actions either *are* those “life” changes, or stand in causal relation to those “life” changes. We might stick to the view that actions are animal changes, that actions are reflexive changes, and so the view that the agent must be this animal. If there is a relation between those animal changes, which are my actions, and distinct lived life changes, the relation will not be one of identity or cause—but rather one of realization. The animal changes, that are my actions, will either realize or be realized by the changes to the lived life. Second, we can, again, set aside the strict reflexivity claim and allow that the actions of an agent are both the reflexive changes, and the concordant changes, to that she is constituted by, or constitutes.

There is, of course, very much more to be said, on the concepts of “life,” “self-conscious life,” “human.” However, note that both those responses take the concept of a “life,” or “self-conscious life” to be the sort of being that is a candidate for being an agent in the first place. And that is far from compulsory. We might rather think of a “life”—a human life, even a self-conscious human life—as involving no more than the human animal living over time exercising her capacities for living. The “life” is what the human animal does when she lives—it is not the agent of what she does.

Perhaps, however, even making this point shows that there remains room for a distinction between the “human animal” and the “alive human animal,” and a room for a dispute about whether the agent is the human animal, or the alive human animal. Is it the “alive human animal” that is the agent that acts, not the human being itself—who can after all also be dead? Maybe. However, this would only stand in the way of the simple solution to our puzzle if we also thought of the “alive human animal” as something distinct from the human animal, while alive. If we do not think that, then to say it is the alive human that acts, is to say that it is the human animal that acts, and our puzzle would still be solved.

7 | A SOLUTION WITHOUT A PUZZLE?

We now get to the second puzzle that I am concerned to explore in this chapter. It seems to be widely allowed among contemporary action theorists that we are human animals, rather than Cartesian, Humean, or Lockean brain-

based selves. Those human animals that we describe as perambulating their way around the world, speaking at conferences on actions, sitting on chairs, eating sandwiches and the rest, for the most part believe that they *are* human animals. There are few Cartesians, Humeans, and perhaps only Parfit thought that he was his brain. (Although Peacocke, see e.g., his Peacocke, 2019, p. 115, also seems to be committed to a related move of making the self the material basis of cognitive operations: the referent of “I” for him is “the material integration apparatus” where the integrating apparatus is the apparatus that integrates the subject’s perceptions and actions.)

Indeed, to be true to our animal natures has been an avowed aim of the most important action theorists we have. Helen Steward is, for example, emphatic about seeking to give an account of animal agency and “to insist that *human beings* are the authors of their actions” (Steward, 2012, p. 21).

This then is our second puzzle: why do we see the three forms of solution to the animal changes puzzle reflected in so much contemporary action theory? We have taken the development of solutions of this form as a cost of distinguishing the self from the animal. Why, then, is that cost also being undertaken by those who do *not* distinguish themselves from individual human animals? What is the need?

I think the explanation starts with the following sort of thinking. What is it for someone to act? Well, it is for them to make themselves move? (see Steward, 2012, p. 15). What do they move? Well, they move themselves, or parts of themselves. How then should we understand the idea that human beings move themselves?

Suppose we start with a general form of a claim to the effect that A moves B, and then think about the case in which an agent moves themselves, perhaps, we will make some progress.

Now, what is the form of a claim to the effect that A moves B? It is natural to take it to involve talk of two things standing in a causal relation. It would also be standard to take it to involve two movings: the moving by A which brings about, or results in the moving by B.

Consider a toy example. Suppose a big rock, that we can call Big Rock, rolls down a slope into a little rock, that we can call Little Rock, causing Little Rock to move. We have in such a situation a case of the form “A moves B,” where there are two objects and two movements such that the movement of Big Rock caused or brought about the movement of Little Rock.

In turn, this kind of familiar tale about objects bumping in to one another, where a movement in one object causes a movement in another object has been taken to correspond to two kinds of movement: transitive and intransitive movement. These labels are taken from a corresponding distinction between transitive and intransitive uses of the verb to move. Big Rock transitively moves Little Rock, and Little Rock intransitively moves as a result. Big Rock in the agent of movement in this situation, and Little Rock the patient. Agents move transitively and patients move intransitively as a result.

It is worth stopping at this point to emphasize that the transitive/intransitive distinction was originally drawn as a grammatical, rather than a metaphysical one. It may be that there is scope for a metaphysical one also, but we should note that the movement between two non-overlapping distinct objects, Big Rock and Little Rock, as described above, stands to be described in two ways.

Consider the movement of Big Rock before it has hit anything. We said nothing about what caused Big Rock to start rolling. Suppose that Big Rock moved because it was hit by Even Bigger Rock. If this had been the case, the movement of Big Rock would both have been a transitive movement—it caused the moving of Little Rock—and would have been an intransitive movement—the moving of Even Bigger Rock caused it. Should we think of this as involving two movements by Big Rock? We would need a very good reason to avoid answering in the negative. The straightforward thing to say is that there was one movement—one rolling—which we take to be an intransitive movement, because it was the result of a movement by another object (Even Bigger Rock), and which we also take to be an transitive movement, because it was the cause of a movement by another object (Little Rock). Moreover, we can happily say that Big Rock’s moving (transitive) Little Rock was because Big Rock moved (intransitive) without thinking that there is a causal relation between two distinct movements.

This distinction between transitive and intransitive bodily movements—corresponding to the grammatical distinction between intransitive and transitive uses of verbs—has, since Hornsby’s seminal discussion in her book

Actions (1980), played a significant role in structuring the contemporary discussion of what an action is. It has been used as the basis of an argument for the claim that what happens when we say that a body moves is a distinct occurrence—in number and in kind—from that which occurs when we say that an agent moves their body. The former is characterized as a *transitive bodily movement*—a BMI, and the latter as a *transitive bodily movement*—BMT. Moreover, when we ask about the relation between these distinct occurrences of different kinds we are invited to take intransitive bodily movements to be somehow the *result* of transitive bodily movements: BMTs are the causes, or causings, of BMIs. I am going to argue, following Haddock (2005) and Lavin (2015), that the grammatical facts do not force a causal distinction of this kind.

Let us return to the case where an agent acts, which we described as an agent moving themselves, or a part of themselves. Suppose, we take as an example, the case of A raising her arm. Following the model we have set out, we might think that this case is also to be analyzed as if it involved two distinct objects involved in two distinct movements—one transitive and one intransitive. We have, on the one hand, A and A's raising of her arm, and, on the other, A's arm, and A's arm rising. Moreover, the earlier model suggested that we think of the relation here as a causal one. A causes her arm to rise. She does so by raising her arm, which is the cause, or causing, of her arm rising. This way of setting things up is evident at various points in Hornsby, Steward, and Hyman and Alvarez (see Hinshelwood, 2017). For them, to say that a self-changing animal is an animal that can make itself move involves:

1. A distinction between an animal that moves and the body that moves—a self-moving animal, capable of agency, *has* a body, rather than *is* a body—and the latter moves when the former moves it.
2. A claim that the animal's action (in its moving its body, or in causing its body to move) and the movement of its body are distinct.
3. A claim that there is a causal relation between the animal's action and the movement of its body: the animal's movement is the transitive movement that is the cause of the intransitive movement of its body (Hornsby), or the animal's action is the causing of the intransitive movement of its body (Steward, Hyman, and Alvarez).

I want to suggest that it may be a mistake to suppose that we can develop a satisfactory account of self-change—*reflexive* change—by starting with a model that involves a causal relation between distinct individuals: an animal and its body. What is critical to the kind of change that actions are is that they are changes involving *one* individual animal, and its relation to itself, and we will not solve the puzzle of agency unless we keep this firmly in view. If that is so then we need to explore the idea that when we say things like “I can make my arm rise, by raising it” or, as Wittgenstein put it “Let us not forget this: when ‘I raise my arm’, my arm goes up.” (Wittgenstein, 1953, 621) the structures that makes true our claim are very different from the structures that make it true that Big Rock's movement was the cause, or causing, of Little Rock's movement. Similarly, the structures that make it true that I moved the chair, and that my movement was the cause, or causing, of the movement of the chair are going to be different from the structures that make it true that I moved myself: I am distinct from the chair but not from myself. If we agree that agency is self-change, and that changes to objects distinct from the agent will be dependent on such self-changes, then we should not be surprised if progress is difficult when we starting our enquires using models on which A moves B is understood as involving a distinct kind of transitive movement by A which stands in a causal relation to a distinct intransitive movement of a distinct object B.

As Doug Lavin, has put it, the way of proceeding makes “the causality involved in action [take] on the appearance of a relation joining particulars, and thus as belonging to the same category as ‘is as large as’ and ‘hates’...it presupposes that the causal element introduced by the transitive verbs employed in ordinary representation of action is a real relation between distinct, fully determinate particulars—some factor *x* and a mere happening” (p. 612, Lavin, 2015).

However, as Lavin emphasizes, nothing in remembering that “when ‘I raise my arm’ my arm goes up,” nor in the ordinary English statements to the effect that a subject can make her arm rise by raising it, necessitates that it involves two individuals and two movements standing in a causal relation. To think that it does is what encourages

the enquiry into which of the elements involved is A's action: is it the rising of the arm, and or the raising of it? Is it BMt or BMi. Or is it, rather, relation between the two—the causing of BMt by BMi?

The point that there is no implication from truths such as “my body moved because I moved it” or “I cause my body to move by moving it” (expressed by employing, transitive and intransitive uses of the verb “to move,” for example) to a commitment to distinct movements of distinct particulars, has also been made by Adrian Haddock, as part of his defense of a view of human actions as identical with bodily movements:

On [Hornsby's] interpretation, “Jane causes her body to move_t” means that the event of Jane's moving_T her body causes the event of Jane's body's moving_i....Hornsby's interpretation tries to display the linguistic truth as forcing upon us a picture of two separate events, the first of which causes the second. But I think we can interpret the linguistic truth in a way that would free us from this picture.

(Haddock, 2005, p. 166)

Suppose one said that the air in a balloon reduces when it shrinks, or that one can reduce the air in a balloon by shrinking it, or “Let us not forget this: when ‘the balloon shrinks’, the air goes out.” All this might be true without it giving one a reason to think that the shrinking of the balloon and the reduction of air in it are distinct occurrences standing in a causal relation. Perhaps the shrinking of a blown-up balloon and the air loss are just the same occurrence—one described in relation to the changing size of the balloon and the other described in terms of the reduction of air. Or suppose one said that the cells in a plant replicate and expand when it grows, or that one can replicate and expand the cells in the plant by growing it. This need not mean that there is an occurrence of the plant growing that is distinct, but causally related to, the occurrence of the cells replicating and expanding: we are free to give an analysis, or constitutive account, of plant growth in terms of cell replication and expansion.

The trouble here, I want to suggest, comes from failing to keep track of the very different nature of the relations involved in reflexive-changes involving one individual and its parts—when A reflexively changes, and the relations involved in non-reflexive changes involving two individuals—when, A changes B. When faced with the latter we will make the following reasonable assumptions:

The causal relation between distinct existences assumption: to explain what A has done—to explain A's causal agency—is to explain the relation between A, and the changes caused to a distinct B.

The distinctness of action and result assumption: given that A and B are distinct the changes caused to B will not be identical to the act of changing B: that will be a distinct changing by A.

If we accept these assumptions when we turn to the case in which A moves herself, we are going to be led to introduce something distinct from A on which she acts, and in relation to which the following assumptions hold.² This is when the appeal to A's body is introduced.

The casual relation between agent and body assumption: to explain what A has done—to explain A's agency—is to explain the relation between A and the changes caused to the position of something distinct from A—her body, or some part of it.

The distinctness of the action and bodily movement assumption: the changes caused to A's body (the bodily movements) cannot be the same occurrence as the act of A changing herself (A moving her body): that must be a distinct changing by A.

This has the result that the bodily movements, that occur when an animal acts, cannot be treated as the actions themselves. The actions, if they are to be found at all, are to be found in the causes, or causings, of these bodily

movements. If the bodily movements—the arm rises, and foot falls—are to amount to actions at all they can only do so in virtue of their relations to other things the agent does.

The explanation for why contemporary theorists, who agree that we are animals, are prepared to bear the costs that we saw being born by the Cartesian, the Humean, and the brain-based Lockean, in relation to the animal changes puzzle, is that they take action to be a causal relation between distinct existences for which the *causal relational between distinct existences*, and the *distinctness of act and result*, assumptions hold. They hold that the power to be a self-changing agent (or “self-moving” agent, in Steward's terms) is not fundamentally different to the power to be a non-reflexive changer: it depends upon *having* a body as a distinct thing that one is able to transact with and make move. It depends on a kind of causal “animal/body” dualism. As Steward puts it:

It is only of some sorts of entity that it makes sense to say that they “have” bodies, thereby separating what is moved (a body or a body part) from what is doing the moving (an animal). It is these entities that are potentially sufficiently complex to sustain an owner/body distinction which I will call “agents,” and the power of self-movement in question is [their] agency...What sorts of entities may be said to have “bodies” to be potentially distinct from them in a way which makes sense of this second stronger conception of self-movement?...Only a creature which can have a mind—i.e. to which certain mental predicates to be applied—can really “have” (own) a body and *vice versa*. There is no point in the distinction between an entity and its body, without the correlative idea of the entity as an initiator, director and discretionary controller of the movements of its body.

(Steward, 2012, p. 17)

Now we should note two things about Steward's kind of “animal/body dualism.” First, it is not any kind of Cartesian dualism—requiring the animal to be a “bipartite construction out of distinct entities” (op. cit. Chapter 1, footnote 37). What kind of dualism it is, is less clear. Second, it is not unmotivated. Steward takes the distinction between self-movers which are capable of agency and *have* bodies, and self-movers (like paramecia) that do not stand in such a “having” relation to their bodies, and lack agency, as allowing us to make sense of the distinction between cases when self-movements are up to the animal—cases in which the animal makes itself move—and cases where they merely “move by themselves” without “making themselves move” (op. cit. p. 17). The commitment is part of an attempt to make sense of the difference between those reflexive changings that are actions by an agent, and those reflexive changings that are not. This is indeed a task that needs to be undertaken, however this way of approaching it, should I think be resisted. It should be resisted for two reasons: first because it gets the structure of the phenomenon to be explained wrong, and second because it does not solve the problem.

It gets the phenomena to be explained wrong, partly because it suggests that our relation to our bodies is instrumental in a way that it seems not to be, and it gives us no sharp way of distinguishing between actions and transactions, even when we think that actions are self-changes, and that there a sharp distinction between agent and non-agent. Moreover, it means we do not get the simple solution to our animal changes puzzle that looks so available and attractive: we cannot say that the animal changes we identify as actions just are the reflexive changes of the animal—they are changes by the animal to a body she has. At best we face something like the position of the constitutivist about human animals discussed in section 6 above.

All this might be worth it if allowed us finally to shelve the problem of how to model *agential* self-change. But it does not. The act of changing her body by A—the causing by A of bodily change—if it is not to be identical to the bodily change that it is a causing of, must still be a change to A. Steward makes a move that needs to be made in appealing to the need for initiation, direction and control. However, if A, as “initiator, director and discretionary controller” initiates, directs or discretionarily controls, a movement of her body then A has changed herself—her initiating and directing are themselves changes to her mind, and changes to her mind are changes to herself. We have no reason, unless we are dualists, for thinking that it will be any easier to explain the self-changes that come with initiations and directions than with arm risings. I want to suggest that in neither case does the agent take herself to be an

object the nature of which she needs to change. In both the agent, herself changes. Whether our *explananda* are cup liftings, and arm risings, or they are initiatings of cup liftings, and arm risings, we are going to have to face the problem of agential self-change square on, and we only move, and complicate, our problem by taking the latter cases, and not the former, as our focus.

The above discussion brings out a feature of the reflexive account. To commit to taking actions to be reflexive changes is not to commit to taking it that the agent, in acting, thinks of themselves as an *object* the nature of which they need to change. It sharply distinguishes cases in which a subject takes an object as an object to be changed—the kettle, or stove, or sandwich, or note—from what is going on when she changes so as to change the kettle, or stove, or sandwich, or note as a result of her agency. Indeed, a mark of our reflexive capacities is that they involve forms of reflexivity in which we do not, and need not, think of ourselves as objects of change. We are subjects of change—subjects who change so that we fulfill our practical aims of changing objects.³ (On this see Guillot & O'Brien, 2022.) There is no doubt that much—perhaps most—of our practical thought is focused, not on ourselves, but on ways the world might best be changed. However, it counts as practical thought only in virtue of the fact that my executing of my decision about what to do—my action—is my self-changing. Much of the action theorizing we have discussed could be happily re-deployed in making sense of the relation between an action and the changes to other objects that the agent succeeds in bringing about in acting. Both the assumption that there is a causal relation between existences (the subject and the object she wants to change), and the assumption that there is a distinction between the subject's action and the resulting changes to the object, do hold true of the relation between the action and the changes to the object an agent seeks to change in changing. However, such deployment will depend on, and not explain, the agent's capacity for action—her capacity to self-change—so that the kettle gets to the stove, and so on.

At the beginning of this piece, I declared an aim to do some ground clearing, and identify where we should *start* our theorizing as action theorists. My suggestion is that we need to work on an account of self-change that does not involve the assumption that self-changes are to be understood in terms of a causal relation between distinct existences: in terms of an agent and the changes she brings about, in something distinct from her, by acts of changing it. To do so will distort the phenomenon, and will leave us without an account of *those* acts of changing. Whether those are supposed to be acts of initiation, direction, tryings, or movings and raisings, which are not intransitive movements and rises, they cannot be explained as acts of self-change in the same way as the bodily acts that we started with without a further bifurcation between agent and patient, and without further bifurcation between the agentive changing and the distinct change brought about. To solve the problem of self-change, we need to work on making sense of how in acting an agent relates to herself, as something that both changes herself, and *in so doing* changes. And actions need to be understood as occurrences that can both be the changing to the self, and the change that is at the same time brought about. To that extent, we do not need to deny the starting idea that in acting an agent (transitively) changes herself, and thereby (intransitively) changes. We just need to resist the idea that that the changing and the change are distinct existences of distinct objects standing in causal relations. Indeed, we may not need to give up on the idea that action involves causing—if we can make sense of causation between identicals, or reflexive causation, or *causa sui*. (Thanks to Tom Smith on this point.)

How to make sense of such reflexive change is a job for another occasion, and the difficulties in doing so are considerable. I end, however, with the observation that, in fact, few action theorists balk, or indeed so much as pause, at the idea that in what they consider to be agential thinking—trying, deciding, judging, directing, calculating—what we have is a single occurrence that is both the thinking by the agent, and thought brought about by the agent. The changing is the change. The task of making sense of this is one that all parties face, and we have no reason to think that explaining the possibility of self-change is any easier—or indeed more difficult—in these “mental action” cases than in the more familiar cases we have been concerned with (see also O'Brien, 2012). Perhaps the power to act, across the board, depends on the possibility of thinking making it itself so. What seems manifest—however we solve our problem—is that trying to make progress in the mental action case by making a distinction between an agent, and some distinct entity that her actions change, say, “the mind she has” and stands in a causal directorial relation to, will get us nowhere. It will duplicate our problem and distort the phenomena.

I have suggested what is manifest in the mental action case is also true of the “bodily action” case, and that for any kind of agency we need to understand reflexive change in such a way that we cleave to the fact that agency involves a relation between an agent and herself: in changing myself, I, myself, change.

ACKNOWLEDGMENTS

My greatest debt is to the editors of this special issue, Alex Geddes and Alec Hinshelwood, for asking me to talk and write on these matters, and for giving very helpful feedback. It is a very good moment in the life of a supervisor when one's students become one's teachers. Sorry to have been a deleterious student at moments. Many thanks to an anonymous referee for this journal for very helpful suggestions. Thanks also to Tom Smith who helped me clarify something that I was worrying over, and getting wrong. Versions of this paper have been presented at the UCL “Being and Doings” conference in 2016, at which some of this set of papers originated, the Bonn “Naturalism, Normativity and the Philosophy of Mind” conference in 2017, “The Lake Geneva Graduate Conference” in Fribourg, 2018, the RIP Manchester “Personhood and Self-Consciousness” conference in 2018, a Humboldt Guest Seminar in Leipzig in 2018, and the ECOM “Kinds of Action” conference 2023. Many thanks to all the audiences—the paper has changed a lot as a result of input.

ORCID

Lucy O'Brien  <https://orcid.org/0000-0001-7569-8739>

ENDNOTES

- ¹ Thanks to the referee for the EJP for raising this issue.
- ² In O'Brien (ms.), “Agency and the First Person” I suggest that the dependence of agency on an assumption of identity of act and result helps us to make sense of our capacity for first person reference.
- ³ They are subjective, not objective, forms of reflexive intentionality, of a kind that echo what Boyle (2024) calls subjective attitudes, and like subjective attitudes the act-types realized by such actions are best characterized in infinitival form—to walk, to wave, to move the kettle, to eat the sandwich, to pass the note.

REFERENCES

- Alvarez, M. (2024). Acting as causing change. <https://doi.org/10.1111/ejop.12944>
- Alvarez, M., & Hyman, J. (1998). Agents and their actions. *Philosophy*, 73(2), 219–245.
- Boyle, M. (2024). *Transparency and reflection*. Oxford University Press.
- Evans, G. (1982). *Varieties of reference*. Clarendon Press.
- Geddes, A. (2016). *Animalism and the human perspective*. PhD University College London.
- Guillot, M., & O'Brien, L. (2022). Self matters. *Ergo*, 9(23), 728–754. <https://doi.org/10.3998/ergo.2617>
- Haddock, A. (2005). At one with our actions, but at two with our bodies: Hornsby's account of action. *Philosophical Explorations*, 8(2), 157–172.
- Hinshelwood, A. (2017). *Thought in motion*. PhD University College London.
- Hornsby, J. (1980). *Actions*. Routledge & Kegan Paul.
- Lavin, D. (2015). Action as a form of temporal unity: On Anscombe's intention. *Canadian Journal of Philosophy*, 45(5–6), 609–629.
- O'Brien, L. (2012). Knowledge of actions and tryings. In A. Coliva (Ed.), *The self and self-knowledge* (pp. 164–179). Oxford University Press.
- O'Brien, L. F. (2017). Actions as prime. *Royal Institute of Philosophy Supplement*, 80, 265–285.
- O'Brien, L. F. (ms.). Agency and the first person.
- Peacocke, C. A. B. (2019). *The primacy of metaphysics*. Oxford University Press.
- Rödl, S. (2007). *Self-consciousness*. Harvard University Press.
- Snowdon, P. (2014). *Persons, animals, ourselves*. Oxford University Press.
- Steward, H. (2012). *A metaphysics for freedom*. Oxford University Press.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Basil Blackwell.

How to cite this article: O'Brien, L. (2024). I, myself, move. *European Journal of Philosophy*, 1–14. <https://doi.org/10.1111/ejop.12944>