Forthcoming in *Philosophical Issues*. Please cite the published version.

AGENCY: LET’S MIND WHAT’S FUNDAMENTAL[[1]](#endnote-1)

Robert H. Wallace

California Polytechnic State University, San Luis Obispo

**Abstract:** The standard event-causal theory of action says that an intentional action is caused in the right way by the right mental states. This view requires reductionism about agency. The causal role of the agent must be nothing over and above the causal contribution of the relevant mental event-causal processes. But commonsense finds this reductive solution to the “agent-mind problem”, the problem of explaining the relationship between agents and the mind, incredible. Where did the agent go? This paper suggests that this challenge against event-causal reductionism is importantly related to debates about fundamentality. It also suggests that extant event-causal answers to the agent-mind problem, ones that suggest that part of an agent’s mind can stand proxy for the agent herself, fail against the challenge. It sketches an alternative reductive view that appeals to entity grounding. This view resolves the commonsense challenge and promises to be theoretically fruitful with respect to other longstanding problems with the event-casual view. The paper concludes with a burden-shifting argument against emergentist agent-causal theories and non-reductive event-causal theories of agency.

1. **Introduction**

The standard theory of action says that an intentional action is caused in the right way by the right mental states (Davidson 1963, Bishop 1989, Schlosser 2019). By filling in the details about causing in the right way and specifying the right mental states, this standard theory can be adapted to explain different kinds of action. Free action, for instance, might be unimpeded intentional action, as Hume suggests (McKenna 2019: 7), or it could be a necessary condition on free action that it is intentional action caused by reasons (Davidson 1973) and their absences (Sartorio 2016: 132).

To secure the right connection between an agent and her actions, this event-causal view must apparently endorse agency reductionism, the view that exercises of agency are reducible to the causal contributions of the appropriate mental event-causal processes (Franklin 2016: 14). Otherwise, it fails to show how an agent brings about (and so controls) her actions. The needed reduction is not eliminative but is instead functional: the role played by the agent must be nothing over and above the right event-causal processes.

Embarrassingly, commonsense finds this reductionism incredible. Can you really believe that the exercise of your agency is nothing more than the causal contribution of whatever set of mental states plays the right role in your mental economy? Don’t *you* cause your actions?Inside the philosophy room, agency reductionism is attractive. It offers a not-so mysterious reductive explanation of the sort philosophers love. Outside of the philosophy room, however, the view seems to say that the agent does nothing. There’s just stuff happening in the mind. I say that viable solutions to the “agent-mind problem”, “the problem…of finding an agent at work amid the workings of the mind”, must pass muster with common sense (Vellemen 1992: 469). And agency reductionism apparently does not.

The force of this challenge to the standard theory is remarkably underestimated, as is, by my lights, the best alternative to the standard theory. Agent-causalism does let you say that *you* are doing the causing.[[2]](#endnote-2) It obviously vindicates the thought that the agent is the one who exercises agency. But the typical version of this view does so by making the agent an irreducible substance cause (Griffith 2016: 73). This irreducibility amounts to saying that, in some sense, agents are metaphysically fundamental. I find that incredible (if flattering). An appeal to metaphysical humility is not an argument. Still, I say that we had better find reasonable reductive metaphysics of agency.

The aim of this paper is to sketch out just such a reasonable metaphysics by employing post-modal explanatory tools. I begin by explicating the agent-mind problem and the event-causal-reductionist view (sec. 2). I then explain the nature of the commonsense challenge to event-causal reductionism and defend its central importance in the context of the theories of fundamentality (sec. 3). I argue against existing responses to this challenge (sec. 4) before advancing a reductive and ground-theoretic answer to the agent-mind problem that diffuses the challenge entirely (sec. 5). This answer promises to be fruitful with respect to two longstanding problems with the standard theory involving alienated agency and deviant causal chains. Finally, adopting a post-modal understanding of reduction shifts the burden of proof. Emergentist accounts of agent-causation and non-reductive versions of the event-causal theory must say more about how they not, ultimately, reductive views of agency too.

1. **The Agent-Mind Problem and Event-Causal Reductionism**

This section explicates the “agent-mind” problem and the standard event-causal reductive solution.

Here’s our problem. As David Vellemen (1992: 469) put it: “the problem of agency is…independent of, though indeed parallel to, the mind-body problem. Just as the mind-body problem is that of finding a mind at work amid the workings of the body, so the problem of agency is that of finding an agent at work amid the workings of the mind.” I prefer the label “agent-mind problem” due to this similarity. Whereas the mind-body problem asks us to explain the relationship between the mind and the body (or mental states and physical states), the agent-mind problem asks us to explain the relationship between the agent and the mind.

What counts as a solution to the agent-mind problem? Chisholm (1978: 622-623) helpfully frames his discussion of the agent-mind problem in terms of “agent-causation”, but not whether there are irreducible agent-causal substances. Instead, he wants to know what explains the truth of ordinary statements concerning agents causing events to happen when they exercise their agency. As Randolph Clarke (2019a: 90) notes, “we may all claim to be agent-causalists” in this ordinary sense. Chisholm’s concern with language is a product of his philosophical context, and it disguises his metaphysical point. We are not talking semantics but “metaphysical semantics” (Sider 2013); a solution to the agent-mind problem tells us the “reality requirements” of thought and language about agents (Williams 2013). Put less linguistically: we want to know whether (ordinary) agent-causation is grounded by anything or if it is fundamental (Schaffer 2009), as the (metaphysical) agent-causal view suggests.

Let’s call the answer the standard theory gives to the agent-mind problem “event-causal reductionism” (or “ECR”). ECR is a combination of two plausible theses:

The Causal Theory: An event is an action if it is caused (in the right way) by the right kind of antecedent mental states.

Agency Reductionism*:* The causal role of the agent is nothing over and above (in some sense) the causal contribution of the relevant set of mental states.

Thus, talk of “agent-causation” is in made true by way of facts about event-causation. An agent exercises her agency and causes her action just when the event of that agent having certain mental states causes her action in the right way.

Each thesis contains a parenthetical that is worth exploring, but only the second one matters presently. (I will return to causing in the right way in section 5). In what sense is the causal role of the agent “nothing over and above” the mental states and processes that cause action? The standard way of defending ECR says that the causal role of the agent is *functionally identical* to the causal contribution of the right mental states and processes (Velleman 1992: 475, Franklin 2017: 582). ECR looks for event-causal processes in the mind that play the functional role of the agent. It reduces exercises of agency as a functional kind to whatever plays this role. Notice too that it would be natural for the defender of ECR to say that ability ascriptions are apt just when the right kind of functional make up is present in an agent. As Markus Schlosser (2019, sec. 2.1) says, “the standard theory of action provides us with a theory of agency, according to which a being has the capacity to act intentionally just in case it has the right functional organization.” More generally, to have agentive modal properties of different sorts—various agentive abilities and capacities— is just to have mental states functionally arranged to play the various roles. Compare with a functionalist story about mental states themselves: pain, for instance, is just whatever state in a system does what pain does, i.e., cause yelling, bring about beliefs about harm, etc. To have the capacity to feel pain is to be put together such that something can play the role of pain.

Functionalism is not always reductive. It is often considered not. Some appeal to the idea that physical states realize mental states. The defender of ECR may appeal to the idea that agent-causation (in Chisholm’s sense) is realized by specific event-causal sequences (Bishop 1989: 96, Clarke 2019a, 2019b). I count that view as reductive with respect to action. Compare a final time with the kind of reductive functionalism endorsed by David Lewis (1983, chap. 9). Given information about context, that we are talking about Earthlings rather than Martians, for instance, we can specify that pain just isC-fibers firing because this firing realizes the functional role of pain in humans. (If we were talking about Martians, pain just is, relative to that context, pneumatic inflations, say). ECR says that, given information about context, namely, what we know about humans as natural beings, the exercise of agency just is certain mental event-causal happenings.

1. **A Challenge That Fundamentally Matters**

The challenge to ECR I will focus on is sometimes called “the problem of the disappearing agent”. It goes like this. If the exercise of agency is nothing over and above certain event-causal processes, one might reasonably ask, “where am I in the event-causal reductionist story?” Nowhere, it seems. The agent has disappeared.

This challenge is intuitively compelling but insufficiently clear. Contemporary discussions are a heterogenous lot: Thomas Nagel (1986: 111) worries about the compatibility of agency and naturalism. Velleman (1992: 462) worries that the standard story fails to explain “the distinctively human feature” of action. E.J. Lowe (2008: 159-161) doesn’t think events have causal power. Jennifer Hornsby (2004: 22) argues that the ordinary role of agents as the cause of what their actions cause cannot be explained in terms of mental events, and Helen Steward (2012: 65-66) takes it that agents settle their bodily actions in a way that renders event-causal analyses circular. A.I. Meldon (1967: 6-9), Christine Korsgaard (2009: 18, 75-76), and Christopher Franklin (2018: 182-188) each argue in different ways that ECR entails that agents are causally excluded from their own agency. What is the common core behind all these complaints?

Perhaps there is no *a priori* route from thought and language about agents to thought and language about events, and so ECR fails. Except this suggestion is question-begging. It assumes an implausibly high standard for successful reduction. As Bishop (1989: 96) and Clarke (2019b: 752-753) suggest, it could be that although agent-causation is conceptually fundamental, it is ontologically non-fundamental such that there is no *a priori* route and yet ECR is true.

My preferred way of framing the problem is admittedly simpleminded. I just think ECR fails the constraints of common sense. As David Lewis (1986b: 134) put it, a theory “cannot gain, and it cannot deserve, credence if it disagrees with too much of what we thought before. And much of what we thought before was just common sense”. ECR seems to disagree greatly with what we thought before, for what we thought before was that agents caused things to happen.

Of course, not everyone likes a commonsense constraint on theory building as much as me and David Lewis. But the naysayers are wrong. Commonsense helps us see what matters with respect to problems of interlevel metaphysics.

I will explain shortly. First, however, we should note some features of our interlevel metaphysics problem, the agent-mind problem. For all their similarities, there is an important *disanalogy* between mind-body problem and the agent-mind problem. The mind-body problem begins and ends with the relationship between physical states and mental states. There is no further question about whether a person’s mind is *their mind* once the reductionist has established that a person’s mind is nothing over and above physical states in a person’s brain. By contrast, the agent-mind problem involves a further question about whether the agent is doing anything once you have established that the agentive modal properties and their exercises are nothing over and above certain event-causal processes. Agents cause things in virtue of exercising their abilities. The agent-mind problem is thus a two-level problem. Indeed, when Velleman (1992: 475-476) articulates ECR, he clearly marks the two levels:

“a person may be an initiator of actions-and hence an agent-in the sense that there is an action-initiating system within him, a system that performs the functions in virtue of which he qualifies as an agent and which are ordinarily attributed to him in that capacity. A reductionist philosophy of action must therefore locate a system of mental events and states that perform the functional role definitive of an agent.”

ECR says that agentive modal properties are to be explained in terms of an event causal action-initiating system. Then there is the further claim that it is in virtue of the possession of this system that someone qualifies as an agent. When an agent acts, the exercise of their agentive modal properties is nothing over and above the causal sequence of events concerning from the relevant system. The direction of explanation runs, ultimately, from the functional organization of a causal system to agents, crossing two levels of explanation.

This feature of the problem—that it involves two levels explanations—appears easily missed, so much so that, as far as I am aware, many of those committed to a version of ECR simply do not say anything explicit about the second explanation at all. This is unsurprising given the sophisticated ways proponents of ECR have dealt with the first explanation. ECR says that causal relations between events of the right sort *just is* agents exercising their abilities. As Laura Ekstrom (2019: 141) puts it: “events—the occurrence of certain of an agent’s attitudes, which are considerations relevant to the decision—cause the event of the decision; that is what it is for an agent to exercise her ability to make a decision for reasons.” And the most plausible versions of ECR say that to have action-initiating systems of various kinds *just is* to have agentive modal properties of various kinds. Indeed, the defender of ECR might insist, as Mark Balaguer (2009: 86) does, that the relevant mental events are “*conscious doing events*, or *conscious choosing events*.” The agent is very much a part of ECR’s explanation of the relationship between underlying functional organization and the nature of agency.

But once you see that a reductive answer to the agent-mind problem requires an explanation *all the way through* from event-causal processes to agentive modality to agents themselves, however, it becomes very easy to unify the disparate expressions of the “problem of the disappearing agent”. ECR needs to say that the agent bears the causal powers of the pertinent event-causal processes but seems to say nothing about this at all (cf. Hornsby 2004: 22). It is simply silent about agents. Thus, the failure to meet the constraints of commonsense is located at the second interlevel explanation between agents and their own agency.

This purported explanatory failure matters for the way we should conceive of fundamentality. Notice that, despite any problems the view may have, agent-causalism (as a metaphysical answer to the agent-mind problem) *handily* answers the second explanatory task by taking causation by agents as a metaphysical primitive.

ECR denies this primitivism, and it patterns with a Humean package view in fundamental metaphysics, one that suggests a fundamental reality composed of combinable simples: “all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another.” (Lewis 1986a: ix). For the Humean, laws of nature, causation, and modality, including agentive modality, reduce to features of this categorical mosaic. As Clarke (1993: 197) says, “the most familiar accounts of event causation are reductionist, aiming to analyze causation in terms of such noncausal and nonnomological features.”

You might not like that ECR pairs with a broadly Humean fundamental metaphysics. I’m not sure I like it. For anyone who want to keep their intuitions metaphysically humble, if not precisely Humean, let ECR stand in as our champion. As I’ll explain presently, alternative reductive views will face similar explanatory burdens. For the commonsense challenge to ECR is an instance of a general challenge to Humean (and otherwise reductive) views. Reductive views have difficulty explaining why things at our level of reality matter. Don’t the more basic constituents of reality do all the explanatory work, such that the less basic constituents simply disappear from our explanations? Conversely, as Sara Bernstein (2021) has convincingly argued, views of the fundamental that take a “mid-level” of reality between the smallest stuff and the whole of everything as fundamental seem poised to “vindicate the reality of social categories, moral properties, and other entities that struggle for ontological respect” within a reductive schema (2021: 1077). “Middleism” arrays with an anti-Humean picture of modality, for middleism can easily explains the causal powers of everyday objects. According to middleist, they are *fundamentalia* (2021: 1077). Agent-causalism is an obviously middleist and anti-Humean view. Many unfairly accuse agent-causalism of making human agents metaphysically exceptional; in reality, the truth of agent-causalism is best seen as part of a non-reductive package deal concerning matters of fundamentality.

So, a Lewisian commonsense constraint is an important methodological consideration with respect to questions of fundamentality, at least for realists about the world of ordinary human experience, anyway. It asks us to recover the world more or less as we found it from the fundament. Humeans and other reductionists generally treat this an important task. Yet (ordinary) agent-causation is a central feature of the world as we find it. If there is no good reductive view of human agency, then we have reason to adopt agent-causalism. And if so, then we may have reason to radically revise our views about the nature and structure of reality writ large. Reductionists about agent-causation— reductionists in general—just haven’t taken the agent-mind problem seriously enough.

1. **Agents by Proxy**

The present question is this: does ECR fail to explain the connection between an agent and her agency, such that we need to revise our thinking about the place of agents in the fundamental scheme of things? Perhaps not. One approach suggests that agent inherits the causal power of the relevant event-causal processes by way of an account of when these processes “can stand proxy for the agent” (Franklin 2018: 191).

A classic example of this kind of proposal is Frankfurt’s (1988, chap 2) view that an agent can lay claim to an action when it is brought about by the higher-order desires with which an agent identifies, for these higher-order desires play the functional role of the agent.[[3]](#endnote-3) Maybe the functional role of agency is played by a more complex kind of desire or conative attitude (Velleman 1992 and 2009, Shoemaker 2003, Jaworski 2007, Sripada 2013), or alternatively by a cognitive attitudes like evaluative judgements (Ekstrom 2000, Watson 2004). Or we could be pluralists and say that many attitudes secure agential approval over the workings of one’s agency (Gorman 2022). In any case, the basic idea is to find an mental system that one cannot “disassociate oneself from” without “forfeiting…one’s identity as an agent” (Watson 2004: 26). Doing so would allow one to properly mark the distinction between events that are “mere happenings” in an agent’s history “and those that are his own activities” (Frankfurt 1988: 59). An agent’s own activities issue from the privileged mental states suitably functionally arranged.

Many of the views just discussed are sometimes labeled as “real self” or “mesh” or “identificationist” views of agency because they suggest that the right relation between an agent and her own agency just is the right relation between various elements of a person’s mental life. Many views offer the agents-by-proxy strategy, however. Consider Fischer and Ravizza’s (1998) view that morally responsible agency occurs when, in the actual sequence of events, a mechanism of action that is responsive to reason brings about an action. Surely, if anything counts as an action-initiating system, such a reasons-responsive mechanism does. Mechanisms of action are modal properties (1998: 53). A mechanism is “nothing over and above the process that leads to the relevant upshot” (1998: 38). Plausibly, this kind of modal property is reducible to the relevant event-causal processes.[[4]](#endnote-4) More streamlined versions of reasons-responsive views make this even more apparent by theorizing directly in terms of the features of actual event-causal sequences (e.g., Sartorio 2016). Reasons-responsive views must also explain how an agent relates to her own agency. Fischer and Ravizza (1998: 211-214) require that responsible agency issues from mechanisms that are “owned” by the agent through a historical process of taking responsibility. To “own” a mechanism, the agent takes up various attitudes towards it, to see the mechanism as an expression of themself as an agent and to judge that they ought to be receptive to reactive attitudes in virtue of the actions the mechanism produces. This is complex instance of the agent-by-proxy strategy.

So, the agents-by-proxy strategy is widespread. But does it succeed against the main challenge? No. To explain, let’s consider a causal exclusion argument aimed at ECR. It’s a dilemma that draws on the work of both Korsgaard (2009) and Franklin (2018).

ECR says something like this: an agent’s exercising her ability to x is nothing over and above the casual contribution of the relevant set of mental states to x-ing. Given this, it seems like the agent could only causally contribute to her x-ing if she was either (1) identical to the relevant set of mental states (Franklin 2018: 182-186) or (2) contributed something over and above the causal contributions of the relevant mental states (Korsgaard 2009: 75-76). But (1) agents are not identical to sets of mental states, and (2) an agent’s exercising her ability to x would not be nothing over and above the casual contribution of the relevant set of mental states to x-ing if the agent contributed something over and above those event-causal processes. So, ECR fails.

The most plausible response for the defender of ECR is to say that agents count as causally contributing to their actions despite not being identical to nor anything above the mental states and events that cause actions (as suggested by Franklin 2018).[[5]](#endnote-5) This is exactly what the agents-by-proxy strategy seems to do. A relevant causal system “stands proxy” for the agent such that the capacities of said system can be attributed to the agent. Franklin (2018: 191-192) understands the agents-by-proxy strategy as offering a non-reductive explanation of how agents relate to their own agency. The strategy thus seems to avoid horn (1) by implicating a kind of inheritance of causal power that parallels non-reductive physicalist views of mental causation (e.g., Pereboom 2011).

Unfortunately, there is a damning problem with this idea.[[6]](#endnote-6) The agents-by-proxy strategy is nothing like the kind of explanation that non-reductive physicalists appeal to with respect to mental causation, which usually involves reference to realized or constituted higher-order properties (Pereboom 2011, chap. 7). In fact, the agents-by-proxy strategy involves reduction, even if its adherents do not mean it to. It appears to be a reductive *paraphrase strategy* with respect to ordinary thought and language.

The agents-by-proxy strategy translates statements about agents into statements about agent-involving (by proxy) event-causation (cf. Lowe 2008: 3-4). Talk about agents doing things is paraphrased into talk about how proxies for the agent, the relevant mental states that play the functional role of the agent, cause events. Talking about the proxies does not require anything of reality that is incompatible with The Causal Theory. Paraphrase allows you to avoid commitment because the reality requirements, so to speak, of the paraphrased sentence differ from the reality requirements of your paraphrase. As Quine (1969: 101) puts it: “paraphrases can enable us to talk very considerably and conveniently about putative objects without footing the ontological bill.” Defenders of ECR do not want to foot the ontological bill of ordinary thought and language, for it seems to contain agents causing things. In paraphrasing away talk of agents causing things to avoid a commitment to agent-causation, however, all that is left are the proxies.

Maybe I’m being too quick. The relevant paraphrases are obviously meant to be reconciliatory rather than revisionary with respect to commonsense (von Solodkoff 2014: 573-757). It is not as if ordinary statements about agents causing thing are *false*. They are true! It is just that they are metaphysically misleading. It is a constraint on a successful paraphrase that it retains the truth value of the original claim, in fact. That there are only proxies left does not show that there is a so-called “problem of the disappearing agent”; talk of proxies is just a more honest way of talking about what happens when agents do things.

Fair enough. Yet paraphrases lack a direction of explanation. For any paraphrase, we might wonder why we should read the paraphrase in an ontologically deflationary way rather than an ontologically inflationary way (Schaffer 2009: 370). Paraphrases, at best, only gesture at ontological dependence. By *not* settling on a direction of explanation to reconcile with commonsense, we render ECR badly ambiguous along both horns of the dilemma. Should we think that we’ve cut away our commitment to agents and so only kept their proxies, or that there really are agents causing things because the proxies are causing things? As Alfred Mele (2003: 231) suggests, the agents-by-proxy strategy leads to “potentially misleading metaphors” rather than genuine metaphysical explanation. It makes it *seem* thatthe agent stands above the event-causal processes and “identifies with” or “owns”.

A final word on proxies. Velleman remarks that his goal is to explain agency “by explaining ow an agent's causal role supervenes on the causal network of events and states” (Velleman 1992: 469). Perhaps, then, the defender of ECR could say that an agent simply supervenes on their proxy, such that you can’t change the facts about the relevant event-causes without changing the facts about the agent.

Unfortunately, like paraphrase, supervenience only gestures at explanatory metaphysical dependence. Supervenience, strictly speaking, is just modal correlation. It’s not an explanatory dependence relation, but rather, a relation that cries out for explanation in terms of explanatory dependence (Kim 1993: 139, Schaffer 2009: 364). Again, we are left wondering if agent-causation (in Chisholm’s sense) has been expunged or if it exists over and above its modal correlate.

1. **A Reductive and Ground-Theoretic Metaphysics of Agency**

In this section, I will argue that we should appeal to grounding, and specifically to entity-grounding, to resolve the main challenge. I will also sketch a solution to some other notable problems for ECR.

Once stated, the form of the solution to the dilemma I raised against ECR is obvious. *Identity* and *distinctness* can come apart. Paradigmatically, they come apart when we consider the relationship between ground and what is grounded (Schaffer 2016: 75–6). For example, a ballet is not identical to several dancers moving thus-and-so, but the dance is not distinct from them either. The dancers *enact* the ballet (Schaffer 2021: 188). It is perfectly commonsensical to say that an agent is not identical to the more basic mental processes that comprise the springs of her action, but it is also perfectly commonsensical that say that the agent is not wholly distinct from them either. We can pass through the horns of the dilemma and meet the constraints of commonsense.

Unlike paraphrase (or supervenience), grounding explanations are paradigmatically asymmetrical. They offer a direction of explanation groundwards from the derivative towards the fundamental (Schaffer 2009: 370). Grounding explanations start at the ground, the fundamental, and “chart how answers percolate upwards through the derivative” (Schaffer 2021: 182). A ground-theoretic version of ECR can take event-causal processes to be the (relative) ground, and percolate explanations towards, and so through to, the agent. The agent is derivative because agency (agentive modal properties and the exercise of agency) is derivative. The existence of agents is grounded in the existence of agency, and the existence of agency is grounded in event-causation by the relevant mental states. By “grounds” here, I mean partial grounds; for an agent, in this context, is also a person. (More on that in a moment).

Talk of grounding is sometimes obscure. Let’s get more precise. We could take a narrow view whereby grounding is only fact-grounding, where one set of facts grounds another set of facts. If so, then we might draw up a solution to agent-mind problem directly in terms of metaphysical semantics. But I think a broad view of grounding as a metaphysical difference-making, such that several relations might count as grounding relations, including realization, constitution, composition, and enactment, is more appealing (Schaffer 2020: 723-724). These relations, after all, are the kinds of things appealed to by non-reductive physicalists about mental causation.

So, a ground-theoretic version of ECR says that agents and agency are realized, constituted, or enacted into being by the right sort of event-causal processes. Constitution turns out to be a useful framework, and as Helen Steward (2012: 55) suggests, ECR seems to say that causal chains of the right sort “simply constitute the settlings of matters by agents”, so that is the kind of view I develop here.[[7]](#endnote-7) The right causal chains constitute the settling of matters by agents.

Many think that constitution is not identity but rather something like non-identity-but-non-distinctness. Puzzles about material constitution can help explain this. David Wiggins (1968: 91) notes that when the lump of clay constitutes the statue, the statue is “nothing over and above” the lump of clay, and yet they are distinct. They exist at the same place and time, but they have distinct modal properties. You could squash the statue and destroy it, but the lump of clay would survive.

Constitution claims, understood as a kind of difference-making entity-grounding, thus have the right form to help ECR. They would allow us to say that agents are not something above the agentive processes that explain action while affirming their ontological and modal distinctness. In fact, it could be constitution all the way down, such that we can plausibly explain how the causal powers of the event-processes percolate up to the agent through agentive modal properties, like abilities and capacities; as Pereboom (2011, chap. 7) convincingly argues in the case of the mind-body problem, constitution claims transfer causal powers from constituents to what is constituted.

One worry with a two-thing view with respect to the statue and clay is “the grounding problem”: if the lump constitutes the statue such that the statue is nothing over and above the clay, then the statue and clay seem to be perfect physical duplicates. One can then ask *in virtue of what* does the statue and clay have different modal properties (Korman 2015: 213-214)? There appear to be no further differences to explain the different properties of the statue and the clay.

Interestingly, the grounding problem shares similarities to two longstanding problems for ECR. For any given proposal about what mental states can stand proxy for the agent, sometimes the agent can be alienated from those states (Franklin 2015 and 2017). More generally, The Causal Theory faces notable problems with deviant causal chains, where the right kinds of mental states cause actions in the wrong way; Davidson (1973) imagines a climber whose beliefs and desires scare him into unintentionally doing what his beliefs and desires rationalized. To wit, it would be wrong to say that agents bring about actions that issue from alienated mental states or that issue from deviant causal processes, and yet these actions do in fact issue from the very kind of functional organizations that defenders of ECR think constitute agency. These problems do not look that different in structure from the grounding problem, which suggests that the statue “disappears” because its existence is fully dependent on the clay; you can destroy the statue without destroying the clay. Likewise, you can “disappear” the agent but not the mechanisms of her agency. In each case, we have reason to think that the intuitive view that there are two things is false. There is just one thing.

The grounding solution to the grounding problem suggests that there are further facts about the statue and the clay that can ground the fact that they are distinct even though they share the same material roots. The natural place is to look at differences in how the parts of the statue and the clay must be arranged. As Daniel Korman (2015: 214) points out, there is a wider range of physical arrangements that can make something clay-shaped than there are physical arrangements that are specifically *this* statue-shaped. And this difference is well placed to explain the modal differences between the statue and the clay. If you smush the statue, it’s underlying constituents will not be statue-shaped anymore but will remain clay-shaped.

My suggestion with respect to alienated agency and deviant action is analogous to the grounding solution to the grounding problem. According to the ground-theoretic conception of ECR, agents and agency share the same metaphysical roots, namely, mental event-causal processes. Yet one pertinent difference between the grounding problem and issues concerning defective agency is that in the former problem, the statue is wholly derived from the clay, but in the case of defective agency, the agent—a person—is not wholly derived from her agency. A person is not just a bundle of mental states or mental event-causal processes. What I should like to say instead is that *whether* the person is an agent, however, is wholly derived, and is a matter of how a person’s psychological constitution overlaps with the constitution of her agency. The relevant event-causal processes can fail to align and become “agent-shaped”, so to speak.

Agents and agency have different arrangement conditions; what arrangements of mental events and processes suffice for their constitution differ. The class of arrangements that suffice for a person to be an agent is much smaller than those that suffice for agency to happen. There’s more than one way to crack an egg, but there’s only one you. For a person to be an agent, there must be significant overlap in these psychological event-causal arrangements, and there usually is.[[8]](#endnote-8) In cases of alienation or deviance, there is insufficient overlap in the arrangement of the event-causal processes for agency to coincide with the rest of a person’s psychology. There is thus no ground for the person to be a well-formed agent with respect to those processes, and so a person cannot rightly lay claim to the workings of her own agency. There is, as it were, a failure of agential constitution.

You might think I’m just pushing the problem around. How could mental events and processes *inside of a persons’ mind* fail to overlap with a person’s psychology? Don’t you need an account of identification or endorsement to explain which parts of a person’s mind get to be “the rest of a person’s psychology”?

No. In trying to locate a “real self” or a particular “mesh” between states or “ownership” over one’s agency, the agents-by-proxy strategy offers the wrong thing, namely a substantive psychological answer to a structural question. We are not looking for a special subset of a person’s psychology that can “speak” for another thing, the agent, as if the agent stands apart from the person who acts. Velleman (1992) was wrong to think that a reductive solution to agent-mind problem requires finding an agent in the mind, as if the agent was already there. Persons *become* agents when things go right. So, mark the distinction between a person and their agency structurally rather than substantively, and in terms of dependence relations, as suggested by my ground-theoretic approach. Alienated agency and deviant action involve *independence* from the rest of a person’s mental life. My ground-theoretic version of ECR says that these cases are failures of agential constitution. Given that the relevant compositional elements are causal relations between mental events and actions, the mark of this independence is a kind of causal insensitivity on the part of the relevant mental events to the rest of a person’s mind. In other words, a person is not able to make a difference with respect to these aberrant causal processes, and so does not fully control them (cf. McKenna 2019).

Of course, this is just a sketch of an answer to these longstanding problems rather than a fully developed response. I won’t offer a fully developed response here. Let me nevertheless sketch on a bit to illustrate. Let’s first consider Davidson’s (1973) climber and deviant causal chains. The climber’s awareness of the danger of holding another climber up by a rope and his desire to survive causes him to loosen his grip out of fear, such that, as Davidson (1973: 153-54) puts it, the climber “never *chose* to loosen his hold”. What is strange about Davidson’s case is the cause of the fear itself. But although this cause is substantively strange, it is not all that structurally strange. Acts driven by fear are paradigmatic instances of failures of agential constitution, for fear is often causally insensitive to a wide swath of a person’s overall psychology. Indeed, the problem with fearful action is usually just that its causal etiology is *too simple.* There cannot be sufficient overlap in the arrangement of the event-causal processes for the fear-driven process to coincide with the rest of a person’s psychology when there are too few elements in the former arrangement. This explains why Davidson’s climber isn’t the agent who intentionally dropped the rope even though he is, of course, the person who dropped the rope.

My view of agential constitution thus helps explain the thought that, as Carolina Sartorio (2016: 137) puts it, “the degree of control a being is capable of achieving is partly a function of the complexity of the causal sources of that being’s behavior.” Yet what matters is not just complexity but also appropriate structure. Given this, I am tempted to think that alienation is not problematic *per se,* but rather, that alienation is normally indicative of a failure of agential constitution. For defective agency is a matter of *bad* agential constitution rather than *absent* agential constitution. Consider a complex case. Say that obsessive compulsive disorder compels me to check to see if the gas stove is turned off. Construe the compulsion as being of sufficient strength and specificity such that it becomes settled that *I just* *hafta* fiddle with the leftmost knob. This check is settled, disturbingly, by some elements of the very functional organization that constitutes my agency; the outsized distress over the possibility of a gas leak gets its grip because I care for the safety of my family. If I am alienated from this checking processes, that would constitute good forensic evidence that there has been a failure of agential constitution, indicating that some of my attitudes cannot make a difference to whether the check occurs. But whether I am alienated from this process or not, whether I desire or value or approve of the check itself—which I sometimes do—or not, the fact that *I am going to check* is not causally sensitive to these mental states, and that is the failure of agential constitution. In fact, the specific process by which I check the knob canbe causally sensitive to what I care about, value, and approve, such that the causal process that constitutes the check reflects significant features my overall psychology, even though I am not a well constituted agent with respect to act of checking itself*.* I might fiddle with the knob while trying to avoid dropping a hot pan on my cat (who is always in the kitchen in the hopes of being fed). I might even temporarily resist the compulsion.

So, being a well put together agent is related to the control someone has over their action, and being well put together is a matter of structure rather than substance. Perhaps, then, we might only be responsible for what we can do around the edges of aberrant agentive processes, for succeeding or failing to adopt protocols aimed at psychic integration, for instance, exposure and responsive prevention therapy aimed at attenuating the causal force of outsized, obsessive distress. Yet not all cases of defective agency are cases where the agent lacks complete control. And we often face obstacles to action involving partially independent event-causal processes, like stubbornly strong preferences, that do not seem exculpating (Lewis 2020: 244). Freedom and responsibility come in degrees. I’d wager that freedom and responsibility are mitigated in accordance with the degree to which an action-initiating system is causally independent from the rest of the person’s psychology. Much more could be said here, but I must pass on the details.

Let me conclude this section by considering a critical objection. Is a ground-theoretic version of ECR *really* reductive? Can I claim the “R” in “ECR”? First, the idea that things could be not-identical-but-not-distinct smacks of metaphysical emergence, as Schaffer himself notes (2021: 201). Some agent-causalists understand their position in terms of emergent agent-causal substances (e.g., O’Connor 2000, Lowe 2008). Second, Clarke (2019a: 78) suggests that “realization may… be available even when a reductive analysis is not.” He, like many, understand realization (and relations like it, including constitution) to be non-reductive.

My view is reductive given a sufficiently liberal construal of reduction. Reduction need not be understood in terms of diminishing ontological commitment by way of identity claims, but rather, could be understood in terms of explanation towards the fundamental. I say this liberal construal is superior. As Kit Fine says (2012: 41-42): “it is only by embracing the concept of a ground as a metaphysical form of explanation in its own right that one can adequately explain how…reduction should be understood.”

Given this, I can shift the burden. Are emergentist accounts of agent-causation *really* non-reductive? They owe a story about how they understand reduction such that purportedly emergent features are truly distinct from the more fundamental stuff from which they emerge. And are realization accounts of agency *really* non-reductive? Realization seems to be a form of entity-grounding. And the reason why ground is important for reduction is because what is derivative is in some sense already present in what is more fundamental. “Grounding relations” are just “ways of separating out aspects that are implicitly present from the start” (Schaffer 2009: 378).

1. **Conclusion**

This paper began with an appeal to metaphysical humility about human agency. In defending a novel version of event-causal reductionism, I hope to have shown that this humility remains reasonable. Yet I have only dealt with one aspect of the larger question: how can we reconcile the commonsense truth that human persons have a metaphysically significant kind of agency with the reductionist thesis that human agency is metaphysically derivative (cf. Nagel 1986: 111-112)? There is still much work to do in pursuit of a satisfying answer.

**Works Cited**

Balaguer, M. (2009). *Free will as an open scientific Problem*. MIT Press. <https://doi-org.ezproxy.lib.calpoly.edu/10.7551/mitpress/9780262013543.001.0001>.

Bernstein, S. (2021). Could a middle level be the most fundamental? *Philosophical Studies*, *178*(4), 1065-1078. <https://doi.org/10.1007/s11098-020-01484-1>.

Bishop, J.  1989. *Natural agency: An essay on the causal theory of action*. Cambridge University Press.

Clarke, R. (2019a). Free will, agent causation, and “disappearing agents.” *Noûs*, *53*(1), 76–96. <https://doi.org/10.1111/nous.1220>.

Clarke, R. (2019b). Agent causation and the phenomenology of agency. *Pacific Philosophical Quarterly*, *100*(3), 747–764. <https://doi.org/10.1111/papq.12275>

Clarke, R. (1993). Towards a credible agent-causal account of free will. *Noûs*, *27*(1), 191–203.

Chisholm, R. (1976). *Person and object: A metaphysical study*. Allen and Unwin.

Chisholm, R. (1978). Comments and replies. *Philosophia*, *7*, 597-636.

Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy* *60*, 685–700.

Davidson, D. (1973). Freedom to act. In T. Honderich (ed.) *Essays on freedom of action* (pp. 137–56). Routledge.

Ekstrom, L. (2019). Toward a plausible event-causal indeterminist account of free will. *Synthese, 196*(1), 127–144. <https://doi.org/10.1007/s11229-016-1143-8>.

Ekstrom, L. (2000). *Free will: A philosophical study*. Westview Press.

Fine, K. (2012). Guide to ground. In F Correia & B Schnieder (eds.), *Metaphysical Grounding*. (pp. 37—80). Cambridge University Press.

Fischer, J.M., and Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge University Press.

Frankfurt, Harry. (1988). *The importance of what we care about: Philosophical essays*. Cambridge University Press.

Franklin, C. (2019). The heart of libertarianism: Fundamentality and the will. *Social Philosophy & Policy*, *36*(1), 72–92. <https://doi.org/10.1017/S0265052519000256>.

Franklin, C. (2018). *A minimal libertarianism: free will and the promise of reduction*. Oxford University Press.

Franklin, C. (2017). Cares, identification, and agency reductionism. *Pacific Philosophical Quarterly*, *98*(1), 577–598. <https://doi.org/10.1111/papq.12204>.

Franklin, C. (2015). Self-determination, self-transformation, and the case of Jean Valjean: a problem for Velleman. *Philosophical Studies*, *172*(10), 2591–2598. <https://doi.org/10.1007/s11098-014-0423-8>.

Gorman, A. (2022). Demystifying the deep self view. *Journal of Moral Philosophy*, *19*(4), 390–414. <https://doi.org/10.1163/17455243-20223564>.

Griffith, M (2016). Agent Causation. In K Timpe & M. Griffith (eds.), *The routledge companion to free will* (pp. 72–85). Routledge.

Hornsby, J. (2004). Agency and actions. *Royal Institute of Philosophy Supplement*, *55*, 1–23. <https://doi.org/10.1017/S1358246100008614>.

Jaworska, A. (2007). Caring and internality. *Philosophy and Phenomenological Research*, *74*(3), 529–568. <https://doi.org/10.1111/j.1933-1592.2007.00039.x>.

Kim, J. (1993). *Supervenience and mind: Selected philosophical essays.* Cambridge University Press.

Korman, D. (2015). *Objects: nothing out of the ordinary*. Oxford University Press.

Korsgaard, C. (2009). *Self-constitution: Agency, identity, and integrity.* Oxford University Press.

Lewis, D. (2020). Outline of “Nihil Obstat: An Analysis of Ability.” *The Monist*, *103*(3), 241–244. <https://doi.org/10.1093/monist/onaa001>.

Lewis, D. (1983). *Philosophical papers, vol. 1*. Oxford.

Lewis, D. (1986a). *Philosophical papers, vol. 2*. Oxford.

Lewis. D. (1986b) *On the plurality of worlds*. Blackwell.

Lowe, E.J. (2008). *Personal agency: The metaphysics of Mind and Action*. Oxford University Press.

Meldon, A.I. (1967). *Free action*. Routledge.

McKenna, M. (2019). “Watsonian Compatibilism.” In J. Coates & N. Tognazzini (eds.) *Oxford Studies in Agency and Responsibility*. Vol. 5. (pp. 5-37). Oxford University Press.

Nagel, T. (1986). *The view from nowhere.* Oxford University Press*.*

O’Connor, T. (2000). *Persons and causes: The etaphysics of free will*. Oxford University Press.

Pereboom, D. (2011). *Consciousness and the prospects of physicalism*. Oxford University Press.

Quine, W.V.O. (1969). Existence and quantification. In *Ontological relativity and other essays*.(pp. 91-113). Columbia University Press.

Sartorio, C. (2016). *Causation and free will*. Oxford University Press.

Sartorio, C. (2014). The problem of determinism and free will is not the problem of determinism and free will. In A. Mele (ed.), *Surrounding free will: Philosophy, psychology, neuroscience* (pp. 255-273). Oxford University Press.

Sider, T. (2013). *Writing the Book of the World*. Oxford: Oxford University Press.

Schaffer, J. (2021). Ground functionalism. In *Oxford Studies in Philosophy of Mind Volume 1*. Oxford University Press. <https://doi.org/10.1093/oso/9780198845850.003.0007>.

Schaffer, J. (2020). Taking causing out of Bennett's Making Things Up, *Inquiry*, *63*(7), 722-744. <https://doi.org/10.1080/0020174X.2019.1656913>.

Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies*, *173*(1), 49–100. <https://doi.org/10.1007/s11098-014-0438-1>.

Schaffer, J. (2009). On what grounds what.’ in D Chalmers, D Manley & R Wasserman(eds.) *Metametaphysics: New essays on the foundations of ontology*. (pp.347-383) Oxford University Press.

Schlosser, M. (2019). Agency. The Stanford Encyclopedia of Philosophy (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/entries/agency/>.

Schlosser, M. (2010). Agency, ownership, and the standard theory. In A. Buckareff, J. Aguilar, & K. Frankish (eds.), *New waves in philosophy of action* (pp. 13-31). Palgrave-Macmillan.

Shoemaker, D. (2003). Caring, identification, and agency. *Ethics* *114*, 88–118.

Sripada, C. (2016). Self-expression: A deep self theory of moral responsibility.” *Philosophical Studies 173*, 1203–32.

Steward, H. (2012). *A metaphysics for freedom*. Oxford University Press.

Velleman, J. D. (2009). *How we get along*. Cambridge: Cambridge University Press.

Velleman, J. D. (1992). What happens when someone Acts? *Mind*, *101*(403), 461–481. <https://doi.org/10.1093/mind/101.403.461>.

von Solodkoff, T. (2014). Paraphrase Strategies in Metaphysics. *Philosophy Compass*, *9*(8), 570–582. <https://doi.org/10.1111/phc3.12150>.

Watson, G. (2004). *Agency and Answerability: Selected Essays*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199272273.001.0001>.

Wiggins, D. (1968). On being in the same place at the same time,” *Philosophical Review*, *77*, 90–95.

Williams, J. R. G. (2013). Requirements on reality. In. F Correia and B Schnieder (eds.) *Metaphysical Grounding*. Cambridge University Press.

1. Thanks to Michael McKenna for comments on a previous draft of this paper. [↑](#endnote-ref-1)
2. See Griffith (2016) for an overview of compelling responses to typical complaints against the view. [↑](#endnote-ref-2)
3. Frankfurt himself was skeptical of The Causal Theory (see 1998, chap 6.), but the strategy under discussion is general. [↑](#endnote-ref-3)
4. Fischer and Ravizza did not explicitly hold the view that mechanisms of action must be understood in terms of The Causal Theory, and Fischer was apparently open to a non-causalist actual-sequence view of free action (See Sartorio 2016: 22-23 for discussion, especially footnote 29). M*echanism* seems like an inherently a causal notion to me, but maybe I’m wrong. [↑](#endnote-ref-4)
5. Franklin (2018: 189-191) notes that we could also reductively identify an agent with some collection of mental events. I think this view is a non-starter but see Franklin’s excellent discussion for actual argument. [↑](#endnote-ref-5)
6. Franklin (2018: 192-194) rejects this solution for a different reason, namely, that it faces counterexamples. I will discuss this problem in section 5. [↑](#endnote-ref-6)
7. See Clarke (2019a) and (2019b) for a realization approach that builds on Bishop (1989). [↑](#endnote-ref-7)
8. This might explain why, as Schlosser (2010: 26) notes, agents can lay claim to their actions by default. [↑](#endnote-ref-8)