



Inquiry

An Interdisciplinary Journal of Philosophy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/sinq20

The embodied, relational self: extending or rejecting the mind?

Joe Gough

To cite this article: Joe Gough (2025) The embodied, relational self: extending or rejecting the mind?, Inquiry, 68:2, 663-695, DOI: [10.1080/0020174X.2022.2075921](https://doi.org/10.1080/0020174X.2022.2075921)

To link to this article: <https://doi.org/10.1080/0020174X.2022.2075921>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 16 May 2022.



Submit your article to this journal [↗](#)



Article views: 1623



View related articles [↗](#)



View Crossmark data [↗](#)

The embodied, relational self: extending or rejecting the mind?

Joe Gough 

Department of Philosophy, University of Sussex, Brighton, UK

ABSTRACT

In putting forward the modern concept of *mind*, Descartes identified the mind with the self. Recently, communitarian and feminist scholars have argued in favor of a conception of the self according to which it includes relations to the social world and parts of the body. If they are correct, it initially seems damning for the view that the self is the mind. I examine whether this is so, by considering whether the identification of self and mind can be saved by recent views of the mind according to which it is partly constituted by aspects of the body other than the brain and by the world beyond the skin. I argue that it cannot, before considering what this means for the mind and its relationship to selves and persons.

ARTICLE HISTORY Received 19 May 2021; Accepted 26 January 2022

KEYWORDS Self; mind; extended mind; relational self; 4E cognition

1. Is the mind the self?

The modern notion of *mind* was posited by Descartes as a self-concept. While the overall person is a union of body and mind, according to Descartes, a person is *essentially* their mind – which is to say, their mind is their self. Several schools of thought, most notably communitarian and feminist scholarship, have recently pushed for greater emphasis on the embodiment and social embeddedness of selves. They emphasize the extent to which people's bodies and social relations (causally) *influence* who they are; they also argue that some aspects of people's bodies and social relations may be *part of* who they are. These points, and their consequences, have been explored and applied in several diverse areas of

CONTACT Joe Gough  jg570@sussex.ac.uk  Department of Philosophy, University of Sussex, Falmer, Brighton BN1 9RH, UK

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

scholarship – bioethics, legal scholarship, political philosophy and traditional philosophical discussions of personal identity, to name but a few.

At first glance, one might think that if these points are correct, then it is damning for the theory that the mind is the self. However, there is a complication. In philosophy of mind and cognitive science, ostensibly similar points have been made about the mind as have been made about the self: there has been a push for the incorporation of more of the body and outside world into the mind. More specifically, it has been argued that the mind is *extended*, partly constituted by elements of the world outside the body, *embodied*, partly constituted by elements of the body outside the brain and *embedded*, depending in important ways on its relations to the outside world. On the surface, these seem like similar points as have been made about the self; it is, therefore, worth examining whether the boundaries of *mind* have been moved so as to line up with the new boundaries of *self*, saving the view that mind and self are one and the same.

In Section 2, I argue that the relational, embodied account of the self is as damning for the identification of self and mind as it first appears, even if the mind is extended, embedded and embodied. While both share a rejection of a certain vision of people as 'bounded', they do not move the boundaries to the same place, in the same way, or for the same reasons. I argue that because of these differences, the mind is not the self, the mind is not part of the self and the self is not part of the mind. More specifically, I argue that on the relational, embodied construal of self, and the extended, embedded, embodied construal of mind, there are parts of the self that are not parts of the mind, and parts of the mind that are not parts of the self.

Having made the central argument of the paper (Section 2), I consider three objections (Sections 3–4). In Section 3, I defend the relational, embodied self against the psychological continuity account and against the animalist account of personal identity. In Section 4, I consider the objection that it is better to try to 'patch up' the notions of mind and self so that mind and self can be identified. In particular, I engage with Taylor's (1989) claim that a mind-like self is a prerequisite for a richer moral language with ideals of autonomy, rationality and rational self-control. In responding to all of these objections, I consider the role that the concept *mind* plays in debates over the self and persons.

2. Disentangling self and mind

In this section, I contrast self and mind (Section 2.4), arguing that the mind has parts the self does not (Section 2.4.1), that the self has parts the mind

does not (Section 2.4.2), and that mind and person come apart in similar ways (Section 2.5). Before beginning my argument in earnest, there are three issues to clarify. Since there have been many different understandings of ‘self’, many of which aim to fulfill wildly different theoretical roles, it is worth beginning by clarifying which understandings of ‘self’, and which such roles, are relevant to this paper – and which are not (Section 2.1). Secondly, it is worth laying out the motivations and minimal commitments of the view of the self as relational (Section 2.2); thirdly, it is worth doing the same for the view that the mind is extended, embedded and embodied (Section 2.3).

2.1. Which self?

There are two notions of self relevant to this paper, and to the majority of the relevant literature. The first – which is what I mean by ‘self’ unless otherwise specified – is something like that person’s ‘real definition’, that is, the sum total of traits that are part of what it is to be that person. This need not be a simple, context-insensitive set of necessary and sufficient conditions on being the relevant person. For one thing, the person’s self may not take the form of necessary and sufficient conditions, but rather the form of a cluster of traits *none* of which is individually necessary, but combinations of which might be *jointly* necessary or sufficient for being the relevant person (e.g. Boyd 1989, 1991, 1999; Wallace 2019a, 2019b).¹ In sum, a person’s self, in this sense, is the collection of that person’s traits which, in the relevant theoretical context, are correctly seen as constitutive of the person, but where constitutiveness does not entail necessity. The second relevant sense of ‘self’, which I mobilize towards the end of this section, is that according to which the self *just is* the person – I refer to this second sense of ‘self’ by directly talking about ‘persons’ or ‘people’.

It is worth stressing some of the differences between *self* and *person* in this scheme. A person is an entity, the bearer of many different properties. These properties may be intrinsic, such as being a certain height, or extrinsic/relational, such as being the child of so-and-so. *Some* of these properties are constitutive of the person, that is, they are part of *what it is to be* that person. The self is this group of properties, those that are constitutive of the person. Roughly speaking, a person (e.g. Mary) has all the

¹Furthermore, depending on one’s interests in discussing the person, different traits of the individual might be emphasised or deemphasised, changing whether or not they are counted as part of the self (eg, Christman 2009).

properties that can be truly predicated of them (e.g. *being tall* if Mary is tall, and *being a mother* if Mary is a mother), whereas Mary's self is the combination of all things that Mary *fundamentally* or *essentially* is. For example, even if Mary is a mother, this does not guarantee that *being a mother* is part of Mary's self, since it does not entail that Mary is *essentially* or *fundamentally* a mother.

One tentative characterization of this difference is that a property Mary has *essentially* or *fundamentally* is one that might *in combination with other properties* be necessary to or sufficient for *being Mary*. It is then an open question, one which might not warrant a single master-theory in response, what grants a trait this status. This question appears often to be settled by trading intuitions about when people in imaginary scenarios would count as 'Mary' (cf. Wilkes 1993). These intuitions, and hence which traits are self-traits if these intuitions are at all reliable, appear to be sensitive to the relative significance of traits in 'shaping' the person – their own view of their life, the way they live their life and their ethically and politically salient characteristics and capacities.

In some recent work on the self, it has been argued that it is necessary to distinguish between two kinds of questions about personal identity and the self (e.g. DeGrazia 2005; Schechtman 2018). 'Traditional' discussions such as that of Locke (1948) and Parfit (1984), are presented as being concerned with numerical identity over time, and it is suggested that views like the narrative self and relational self deal with a *different* and hitherto-neglected issue, namely, that of the proper *characterization* of a person. To some extent, the idea seems to be to concede that 'traditional' discussions of the self are on the right track, and not challenged by these more recent accounts of the self, while carving out 'new' terrain for these more recent accounts to occupy. I will not be adopting this strategy here, for two main reasons. First, it seems to me to concede too much, since relational accounts of the self can hold their own on traditional terrain. For example, although the relational self is often put forward as a revision of standard concepts of the self, empirical works suggests that 'children [consider] moral traits connected to interpersonal relationships crucial for preserving personal identity' (Jirout Košová et al. 2021, 47). Secondly, and moreover, I am not convinced that the numerical identity question and the characterization question can or should (let alone must) be separated. According to my view of persons, they are 'natural individuals' (Boyd 1999), with real definitions that take responsibility for *both* their numerical identity *and* their best characterization. As such, I treat the relational account, and the more 'traditional' animalist and

psychological-continuity accounts as genuine competitors, ultimately offering different accounts of the same thing.

This understanding of *self* and *person* resonates with much of the literature on selfhood and personal identity. It does not, however, resonate with notions of self according to which the self is a substance, an object or an entity *that is distinct from the whole person*. Such notions include the notion of a *minimal* self, apparently motivated in order to account for self-awareness, especially the kind of pre-reflective self-awareness supposedly essential to subjectivity (e.g. Zahavi 2008), and some notions of self apparently motivated by the need to account for self-reference. These notions of self are not relevant, at least not obviously so, to the argument of this paper. However, in the interests of transparency, I should note my suspicion that these notions of self are unmotivated, if they are supposed to be distinct from the whole person: I suspect that the 'selves' of self-reference and self-knowledge are just *oneself*, that is, the whole person, and hence these notions of self must either be reduced to the notion of a whole person or eliminated (e.g. Strawson 1959). It has a better, but imperfect, fit with notions of self according which effectively reduce one's self to one's *sense of self* – one's sense of self may be a good indicator of which traits are parts of one's self, and it may also be a mechanism by which traits come to be part of one's self, but the real definition of a person may well include or exclude traits that are part of that person's sense of self.

2.2. The relational self

In sum, a person's self is the real definition of that person, the collection of traits that make them the particular person that they are. The notion of a *relational* self can partly be understood as a result of feminist theorists' attempts to reconceive the self in a way that supports the notion of relational autonomy, often contrasted with a traditional 'liberal' notion of autonomy (see especially Mackenzie and Stoljar 2000; Nedelsky 2011). Such a notion of autonomy, and the associated understanding of the self, is accused of encouraging, or perhaps embodying, the atomistic, individualistic and masculine-coded attempt to strive for independence from others, while downplaying the role and needs of the feminine-coded body as well as feminine-coded emotionality.² Thus, the relational self

²It is important to stress here that the dualism between body and mind, as well as the idea that the body and emotion are bad and feminine and that the rational mind is good and masculine are not *endorsed* by feminist scholars. The gendered dualism and associated evaluations are offered as a description of

is a notion of self that is geared towards an understanding of human freedom, and human wellbeing, according to which the influence of social relationships, emotion and the body may be both beneficial and extremely significant for people. It is worth noting, however, that just as 'relational autonomy' is the name for a family of theories of autonomy that set out to achieve these aims (e.g. Mackenzie and Stoljar 2000), so the 'relational self' is the name for a family of theories of the self.

In both communitarian and feminist scholarship, the notion of a relational self can be seen as an attempt to resist the idea that people are detachable from their communities, cultures, and social lives (e.g. Taylor 1985). These communitarian critiques are *also* tied to an emphasis on embodiment (e.g. Taylor 1999), albeit for slightly different reasons. A further point worth noting is that these concerns go back further than contemporary feminist and communitarian scholarship, the latter of which can largely be understood as resisting Rawlsian liberalism (Rawls 2005), at least to Levinas' (a contemporary of Hitler) worries that the disembodied, detached liberal self-conception laid the groundwork for the rise of Hitler's politics (Levinas and Hand 1990). Similar concerns are also plausibly found earlier still, in the Hegelian and Marxist critiques of Kant and Fichte's notions of freedom, accused of 'smuggling in' the norms and institutions of private property (e.g. Rose 1981, 59–62).

Often, the considerations adduced in favor of emphasizing the extent to which people are embedded in their social worlds are claims related to the development of desirable capacities – freedom, wellbeing,³ rationality, authenticity and so forth (e.g. Taylor 1985, 1989). The idea is that for people in general, much of the value of human life relies on being embedded in social life, and that for individual people, their social situation is much of what makes their life the kind of life it is. It would be absurd to deny that the liberal understanding of selves is directly committed to the rejection of these causal-developmental claims (e.g. Christman 2009). For example, Mill – a liberal if ever there was one – endorses similar claims.

In discussing the human sciences, Mill suggests that alongside associationist psychology (discovering the lawlike regularities of conscious life), there needs to be a science of ethology. This science would discover

the conceptual framework these scholars are inheriting and criticizing. Some respond by challenging the idea that the body and emotion are bad, others by challenging the idea that they are feminine, others by questioning the dualism, and others still by challenging some combination of the above.

³An off-the-shelf theory of wellbeing which resonates well with such relational accounts, and has the virtue of being entirely independently motivated, is Bishop's (2015) network theory of wellbeing.

the kinds of social conditions and education that (along with human nature) put the psychological laws in place, and ultimately explained them. This scheme suggests that Mill sees people's inner lives as ultimately shaped by the social conditions which, along with inherited human nature, put the 'laws' in place (Mill 1843/1974, see especially ch. 6).

As such, if a theory that merely avows these causal-developmental claims counts as offering a relational self, it marks no genuine contrast with traditional liberal notions of the self. On the assumption that the relational self marks a genuine contrast with prior, 'liberal' understandings of the self, the relational self must go beyond these causal-developmental claims. A plausible alternative is to claim that the liberal understanding of selves does not give sufficient *emphasis* to these causal-developmental claims, because it is ultimately committed to an understanding of selves that precludes social conditions and relationships from being part of one's self. On this view of the debate, theories of the self as relational assert, and their opponents deny, that these (causally) significant social relationships and conditions might actually be part of what it is to be who one is. What makes a theory of the self 'relational' is the claim that for many people, social conditions and relationships are *part of* their selves. Likewise, what makes a self 'embodied' is bodily features being part of the self.

2.3. The 3E mind

On, now, to the view of the mind ostensibly serving as a natural corollary to this view of the self. The extended, embedded and embodied view of the mind all deny that our explanations of mental phenomena should focus on what's going on inside our skulls. The extended and embedded views focus on relations of causal coupling – dense, reciprocal and causal interactions; both focus on the coupling between the brain and the world outside the skin, although they draw slightly different conclusions from this (e.g. Clark 2013; Wilson and Foglia 2017). To claim that the mind is extended is to claim that the coupled parts of the world are, at least sometimes, part of the mind; to claim that the mind is embedded is to claim that the mind's functioning is importantly dependent upon this coupling, without going so far as to claim that the coupled parts of the world are part of the mind. There is an ambiguity in claiming that the mind is embodied: it can be taken merely as the claim that the body (beyond the brain) causally influences the mind, but is not part of it; it can also be taken as the view that there are parts of the body that are part of the mind. This

latter interpretation of the claim is the one relevant here, and it again relies on a claim of coupling, but here between the brain and other parts of the body (Varela, Thompson, and Rosch 2016; Wilson and Foglia 2017).

The history of these ideas, or at least the history of their most recent incarnation, is well understood. The view of the mind as embedded grew out of the *situated cognition* trend in cognitive psychology (e.g. Danziger 2008), and the view of the mind as extended grew out of the Anglo-American philosophical tradition and functionalist approaches to the mind (Clark and Chalmers 1998), and the view of the mind as embodied grew out of the phenomenological tradition, especially the work of Merleau-Ponty, being put to work in cognitive science (e.g. Varela, Thompson, and Rosch 2016; Wilson and Foglia 2017).

2.4. *Self and mind*

All of these views of the mind are primarily motivated by a wish to move towards a full and accurate scientific understanding of human cognition. This is the first contrast to be made with the view of the self as relational and embodied – the motivations for that view are primarily ethical and political. Another contrast is that while work on the relational self questions the very idea of the self having boundaries (e.g. Nedelsky 2011), the extended and embodied mind views largely retain the idea (associated with their main opponent, the neurocentric view of the mind) that the mind is a concrete, bounded, material system and argue that more material is part of that system than previously acknowledged. A third key contrast is that different senses of ‘parthood’ are used in the discussions of mind and the discussions of self. *Being part of the mind* is a matter of being part of the material system that implements mental processes in the above claims. Conversely, *being part of the self* is a matter of featuring in the real definition of the relevant person.

It is important to note, however, that it is not as simple as the relevant self-talk being interested in ethics and essences, while the relevant mind-talk is interested in scientific explanations and causation. This contrast disintegrates in large part because claims about the (causal) development and inculturation of ethically salient capacities are vital to arguments for the relational self. This comes out most clearly by considering theories of what makes a trait part of the self conceives of people as a network of traits, and suggests that those traits which are part of the self are those which are strongly relevant to the overall organization of the network,

that is, those traits that play an important organizing role for the overall network. Wallace's relevance relations are neutral between entailment, parthood, and causation. For example, a person's *being a professor of philosophy* might be a strongly relevant trait, and therefore a part of their self, because it has an important organizing role, entailing many traits (e.g. that one has spent a long time in education) causing many other traits (perhaps arrogance, a certain view of one's self and a probable life-course), and being constituted of several other traits (e.g. being a professor, studying philosophy and being employed by a university). That a person is a professor, in other words, tells us a lot about what it is to be that person, and therefore is part of what it is to be that person. One significant feature of Wallace's view which I find appealing is that it can account for the significance of one's *sense of self* without reducing the self to the mere sense of self – one's sense of self provides an important mechanism by which traits might play an organizing role.

Another theory can be suggested by drawing on the work of Mead (1934). Mead offers a particular theory of people's cognitive and social development, according to which one develops one's own perspective on the world by developing the ability to reflectively and reflexively engage in a sort of 'internal conversation' between different perspectives that one has internalized from others. Pride of place in this theory is afforded to a kind of socially-engaged cognition that has generally fallen out of social psychology since the early twentieth century (Greenwood 2004). This is socially-engaged cognition where one thinks what one thinks on the basis of group membership – one adopts attitudes and opinions on the basis of recognition of one's group-membership, and the perceived attitudes and opinions of the group of which one is a member. According to Mead's theory, one develops one's own perspective and attitudes by putting the attitudes and opinions one has received from one's social environment (from influential individuals in one's life such as one's parents, and from this kind of socially-engaged cognition) into conversation with one another within oneself. One's relationships with these individuals (e.g. *being a child of...*) and groups (*being a member of...*) may be part of what it is to be the person that you are, by being incorporated in this way into one's own perspective on the world.

Yet another theory is suggested by the arguments from the causal-developmental claims to the nature of the self. In order for these arguments to be valid, some kind of principle is required according to which those traits which are significant preconditions for the

development of a person's most valued or valuable characteristics, are part of that person. It is worth noting that these theories are not necessarily in competition in any interesting sense: they may all be understood as providing principled reasons for counting a certain trait of a person as a part of themselves.

None of these theories is uninterested in the causal workings of the world. Each makes claims that can be taken as scientific hypotheses and the relationships (causal and otherwise) between various traits, claims that might be accepted or rejected on purely non-ethical grounds. Further undermining any simplistic contrast between the workings of these accounts of self and these accounts of mind, there has recently been a movement in discussions of the extended mind towards using ethical considerations as part of the criteria for individuating the material systems that underlie mental processes (e.g. Vold 2018).

2.4.1. *Self beyond mind*

Nevertheless, if selves are relational and embodied, and minds are extended, embedded and embodied, self and mind come apart. One important issue I have thus far glossed over is whether parts of the outside world or merely one's relations to those parts of the outside world, might be part of the self. Consider, for example, a person for whom *being the child of Steve* is one of their defining traits, and therefore part of their self. Is *Steve* a part of that person's self, or merely *being a child of Steve*?

Either way, the self exceeds the mind. Imagine that *Steve* died before the person was born. If *Steve* is a part of the person's self (in my view, the less plausible option),⁴ then there is clearly a part of that person's self that is not a part of their mind. The reason for this is that the person is not in a dense reciprocal causal relationship with *Steve* (who is, by stipulation, dead), although *Steve* is nevertheless part of the person's self.

If it is merely *being a child of Steve* that is part of the person's self, it is still highly likely that this would not be a part of their mind. One reason for this is that it is simply not clear what it is for a relation like this (let alone a social role like *being a professor*) to be a part of one's mind, where this is understood as a matter of being part of the material system that

⁴I find this less intuitively plausible, but I also think that the debate over relational notions of autonomy pushes us in this direction. In order to avoid 'perfectionism', whereby the standards of autonomy are too stringent if social conditions are directly built into the notion (Christman 2009), while retaining the benefits of a relational account of autonomy, the solution appears to be to focus directly on the (kinds of) relations required for autonomy (Westlund 2009). In as much as the relational self is motivated to account for relational autonomy, similar considerations apply.

implements mental processes and capacities. Furthermore, even if one can make sense of the idea of such a trait being part of one's mind, it is not something that can be causally affected by the person in question, precluding the possibility of reciprocal causal interaction of the sort emphasized in the relevant views of the mind.

For similar reasons, the relations that are part of the relational self are also not those that 'embed' the mind – the relations that are part of the relational self need not be causal, let alone dense or reciprocal. These relations, therefore, need look nothing like the 'coupling' characteristic of those relations emphasized by the embedded view of the mind.

More generally, the emphasis in discussions of the relational self lies squarely on the influence *on* the person *of* (their relations to) the outside world (and their body) – as reflected by the emphasis on development and inculturation. This runs counter to the extended, embedded and embodied mind views' emphasis on coupling, which is a matter of dense, and *reciprocal* interaction and influence. Similar points apply to the kind of bodily traits that might be part of the self. *Being tall* (to choose a benign and banal example) might well be part of someone's self, but it is not something that they are likely to have had much impact upon (and again, is something which it hardly makes sense to claim is part of the material system underlying mental phenomena), and so would not be counted a part of their mind.

2.4.2. *Mind beyond self*

So far, I have argued that the self has parts that the mind does not – and, therefore, that the self is not part of the mind (if such a thing even makes sense). I now want to focus on arguing that the mind, if it is extended, embedded and embodied, has parts that the self does not – and so that the mind is not part of the self. Traits generally get counted as part of a person's self in large part because of their overall significance in shaping that person, their perspective and the course of their life. Just as the focus on coupling precludes certain parts of the self from being parts of the mind, so this focus on depth of influence precludes certain parts of the mind from being parts of the self.

Consider, for example, a student doing long division with pen and paper, where this is partly implemented by the pen and the paper used; consider also the bartender, who remembers people's orders by arranging and rearranging glasses along the bar, thus offloading some of the work of memory onto the outside world (Clark 2013). If coupling relations were all that is required for being part of the mind, then these

would be plausible cases of extension. These parts of the outside world (and one's relations to them) are not promising candidates for being part of the self. The reason for this is that although they have a significant short-term causal-explanatory role with respect to a particular cognitive capacity, they may well have no significant role in shaping the student or the bartender – their life, perspective or ethically significant traits. Furthermore, and setting aside extension, embodiment and embedding, it seems highly plausible that many mental traits are not part of the self. *Being able to do long division* and *being able to remember people's orders well* are fairly uncontroversial examples of mental capacities, but this does not give us any good reason to believe that they are at all significant in the person's life, let alone part of the person's self.

According to most formulations of the extended mind, however, coupling alone is *not* sufficient for being part of the mind. To avoid (a perceived risk of) absurdity, further criteria have been added, most notably, that the external component be readily available and reliably accessed, and that any information contained therein should be automatically endorsed, be easily accessible and have been consciously endorsed (Clark 2013, 197). I use notebooks a great deal, which meet these criteria, mostly for shopping lists and reminders. These notebooks and my relations to them, however, are not (I hope) part of myself – they simply are not that important to me, and I would barely be different without them. The point here is that these further criteria do not ensure the kind of significance required for being part of the self, although it is worth noting that they may well bring the extended mind into *slightly* closer alignment with the relational self by ruling out the most transitory cases of coupling.

The upshot of all this is that the relational, embodied self and the relational, extended and embodied mind each have parts the other does not. As such, using these 'off the shelf' theories of self and mind, the self is not the mind, and neither is part of the other. All this, so far, has been framed using the understanding of 'self' according to which the self is the real definition of the person. I now want more briefly to consider the second relevant sense of self I outlined at the beginning of this section, according to which the self is the person.

2.5. Mind and person

Before discussing the relationship between the mind and the person, it is worth fleshing out the claim that people are relational. The claim that

people are relational needs to go beyond the claim that people *have* relational properties, and beyond the claim that they are *influenced* by relational properties. The best way to flesh out this view, in my opinion, takes its cues from a debate in metaphysics over the nature of change. Geach (1969) objects to the ‘Cambridge’ (i.e. Russellian) account of change according to which to change is to have a property at one time and lack it at another on the basis that this picks out far too broad a category of ‘changes’.

The problem, Geach argues, is that a change in one’s relational, or extrinsic, properties is not a *genuine* change – for example, I do not (genuinely) change if a tree falls in Morocco, but by the Cambridge account, this is just as much a change as if my heart explodes. Geach proposes that *genuine* changes are limited to changes in one’s *intrinsic* properties. This account of change is unduly restrictive: as Weberman (1999) argues, there is an important class of entities for which a relational/extrinsic change is nevertheless a genuine change – for example, a £10 note locked away in a perfectly sealed safe genuinely changes (in value) if civilization collapses and the world becomes some kind of post-apocalyptic dystopia, *regardless* of whether any intrinsic changes in that £10 note come about as a result.

The relational account of persons should, I think, be the claim that persons fall under this class of entities, for which a relational change can be a genuine change. Note that the class of properties which may be involved in a genuine change may be broader than the class of properties which are constitutive of the person, and, therefore, part of the self: a genuine change need not be a change in a self-property. For example, a £10 note might change in its precise exchange-value without any of its constitutive properties changing.

Interestingly, one example Geach mobilizes is that of *becoming a widow*, which he counts as a non-genuine change, claiming that the only genuine changes associated with *becoming a widow* are the resultant intrinsic changes. In opposition to Geach, I think that even if *no* intrinsic changes come about as a result of one’s becoming a widow, this might (although need not) constitute a genuine change in the person who has become a widow. The conflict between these intuitions on this example, I think, neatly captures what is at stake between a relational and a non-relational view of persons. I find the view that *becoming a widow* could count as a genuine change in the absence of intrinsic changes hardest to deny when considering a scenario where, although the widow does not *know* whether her spouse is alive, she cares

whether they are and believes that they are. Imagine, for example, that she regularly prays for her (dead) spouse: whether praying for one's absent spouse is touching or tragic⁵ depends in part on whether that spouse is alive. The tragedy of losing a spouse (about whom one cares) is surely not entirely eliminated simply by never finding out about it.

On, now, to the question of the relationship between persons and minds: clearly, the person is not part of the mind – the person has parts (e.g. the appendix) and properties (e.g. a certain height) that the mind does not. However, it seems eminently intuitive that the mind is part of the person. It is my view, however, that defenders of the extended mind view (in particular) ought to deny this claim. I believe that they ought to take their arguments for the extended mind *not* as showing that parts of the world outside the body are part of the person in virtue of being part of the mind, but *rather* as showing that the mind is, in fact, not a part of the person (although it may nevertheless belong to the person, and hence people may nevertheless *have a mind*).

Such a view has the disadvantage of denying the intuitive claim that the mind is a part of the person. But some level of counter-intuitiveness is a general problem for the extended mind view, since either way, the account is committed to at least one counterintuitive claim: either the denial that the mind is a part of the person or the assertion that certain parts of a person's environment are parts of the person. The advantage of denying that the mind is part of the person, over the alternative, is that this fits better with other apparent truths about people – for example, that when faced with a person outside of their normal environment (and without those parts of their normal environment that might be counted as long-standing parts of their mind), we are faced with the *whole* person. When someone sees me without the memory aids to which my brain is plausibly coupled, and which I reliably and frequently access, (that is, without my many notebooks), they are seeing *me*, not just *part of me*.

One might worry that this loses some of the ethical appeals of the extended mind. For example, there is a class of Alzheimer's sufferers who remain high-functioning in their home environment, despite performing poorly on standard cognitive tests, because they have offloaded a significant chunk of their mental function to their home environment (e.g. Clark 2003; Drayson and Clark, [forthcoming](#); Vold

⁵I hope it is clear that this is not the 'merely subjective' issue of how people are disposed to react, but rather the issue of how people *ought* to react, and what reaction the unknowing widow and her life *warrant*.

2018). It is supposed to be an ethical benefit of the extended mind that it captures what is so wrong with forcibly removing such a person from their home environment – to do so is, on the extended mind's account, to deprive them of their mind, and therefore to deprive them of a part of themselves. My response is that one can capture the problem with removing a person's mind without claiming it is part of them – a person's mind might be a terrible thing to take away from them, even if it is only something they *have*.

One might also claim, in the vein of the relational self (not taking 'self' to mean person, but in the default sense), that these people's *relations* to their home environments are part of their selves, because of the huge role that these relations play in shaping their lives. More generally, some of the relations one has to one's environment might meet the criteria *both* for being part of the self, *and* for extending the mind, that is, there might be coupling relations of the right kind for extending the mind that *also* have great significance in shaping the person and their life. This kind of significance and this kind of coupling do not *preclude* each other; but they also come nowhere near *guaranteeing* each other's presence.

To recap, I was examining whether the extended, embodied, embedded mind proposal can save the identification of self and mind from the criticisms of those calling for a relational, embodied account of selves and a relational account of persons. I have argued that it cannot, because, despite superficially similar points being made about minds, selves and persons, these views entail that the mind is not identical to the self or the person, that the mind is not part of the self or the person and that the self (and person) are not part of the mind. The different emphases in discussions of mind and self lead to radically different criteria being adopted for what it takes to be part of the mind, and what it takes to be part of the self. While these views reject a narrowly bounded view of persons and their make-up, they move the relevant boundaries to different places, in different ways and for different reasons.

3. Objections from the 'big two'

At this point, I have argued that if the self is relational, then it is not the same as the mind even if the mind is extended, embedded and embodied. For some, this argument may seem wrongheaded or pointless: many people would reject the antecedent, and may therefore have no interest in the consequences of taking a relational view of the self. Two particularly popular alternative accounts of the self (in the relevant

sense, where the self is the real definition of a person) are the psychological continuity account and the animalist account. If either of these accounts is correct (at least on any standard construal), then the relational account is false.

There are many variations of the psychological continuity account. It would neither serve my argument nor do justice to these accounts to survey all of them here. As the name suggests, the idea is that some kind of psychological continuity accounts for a person's persistence over time. Often, memory is emphasized in such accounts (Locke 1689; cf. Behan 1979), although some versions of the psychological continuity account adopt a more general notion of psychological continuity (Olson 2021; Shoemaker and Swinburne 1984). What these accounts share, importantly for my argument, is the view that all of the traits that are part of one's self – all of the traits that make one who one is – are mental traits.

Animalist accounts come in two main forms: organicist and somaticist. Organicist animalists, for whom life is essential to identity, hold that the self comprises those life processes involved in 'maintain[ing] the organism's complex internal structure' (Olson 2007, 28; Blatti 2014). Somaticist animalists, for whom life is not essential for identity, hold that the self comprises the (spatial) organization of our (physiological/anatomical) constituent parts (e.g. Mackie 1999; Blatti 2014).

My response to this kind of worry is threefold. The first part is to note that, even if correct, only very few psychological continuity accounts would save the identity of self and mind. Animalist accounts clearly entail that the mind is distinct from the self, since they reject the idea that *any* mental features are part of the self. Psychological continuity accounts all share the idea that the self has only mental parts – and so that the self is part of the mind. According to most versions, the mind has parts that the self does not – not all of one's mental states, features or capacities are relevant to determining a person's persistence conditions. Shoemaker's functionalist account of personal identity seems perhaps to save the identity of self and mind – every part of the mind, on his account, is at least potentially relevant to the persistence of the person, albeit primarily because he believes that all mental states are defined in relation to one another (alongside perceptual inputs and behavioral outputs).

The second part of my response is to highlight that each sort of account faces major difficulties – many of which have been explained in great depth by supporters of the opposing account. I find the

objections offered by psychological continuity theorists to animalist accounts extremely compelling. I share the strong intuition that if *all* of Sam's psychological properties were transferred to Steve's body, then Steve's body would become the home of Sam's person. Even if this is wrong (e.g. for the reasons suggested in Wallace 2019a), the intuitive weight of this example seems to me to be strong evidence that psychological traits are at least *somewhat* relevant to identity-determination and, therefore, part of the self. I will explain further below why I do not believe that, even if this intuition is correct, it proves the psychological continuity account.

I do not find psychological continuity accounts compelling because I believe that there are cases of identity-persistence without psychological continuity. As such, even if *some* people's 'real definition' comprises solely mental traits, it is not plausible that this is the case for *all* people. Lindemann (2009) discusses a particularly compelling class of cases for which the psychological continuity account is insufficient, where the family of a sufferer of dementia may 'hold' that person in their identity primarily by 'treating [them] in accordance with their narrative sense of [them]'. In such cases, although the sufferer of dementia no longer has many of the capacities taken as essential to personal identity by psychological continuity accounts, they are held in place as the person that they are by their relations to and interactions with those around them.

The final part of my response does not target the accounts themselves, but instead the arguments for these accounts. Often implicitly, sometimes admirably explicitly (Shoemaker 2011), the debate between animalist accounts and psychological accounts proceeds on an *exclusivity assumption*: either all the traits that are part of the self are mental, or none of them are. Additionally, it is assumed that those that are non-mental are biological (either life processes or spatially arranged physiology/anatomy). These assumptions are also highlighted in Wallace (2019a, 2019b).

There are at least two problems with these assumptions. The first is that a large class of the properties that we have and that are significant to us – relational and extrinsic properties – is written off as accidental to who we are prior to the debate. The second is that 'mixed' accounts are ruled out prior to debate – accounts according to which both mental and non-mental features contribute to identity-determination.

Without these assumptions, the argument from mind/brain-transfer cases to the psychological continuity account fails. Suppose, for the sake of argument, that if all of Sam's mental properties are transferred from Sam's body to Steve's, then Sam moves to Steve's body. This

clearly shows that mental properties are part of Sam's self. However, it does not show that *only* Sam's mental properties are part of Sam's self. Instead, it shows that *enough* of the traits that are part of Sam's self are mental that transferring *all* the mental properties results in identity-transfer. It might also be the case that transferring half of Sam's mental properties, and some of Sam's non-mental properties, suffices for identity-transfer (while transferring only those mental properties does not suffice).

Likewise, the argument for animalist accounts is on a shaky footing without the exclusivity assumption. Animalist accounts are motivated by the claim that (human) people are humans, and humans are animals – so far, so good. Why exclude mental properties from those involved in the individuation of animals? If the motivation of animalists is to defer to biology, this is a strange assumption, especially for organicist animalists. There is no reason to think that biologists or lay folk, as a rule, ignore psychological (or indeed relational) traits when individuating living systems (e.g. Boyd 1999) – it seems, at best, a hangover of a dualistic view whereby the mind is beyond the scope of biology and the scientific view of the world. If organicists wish to account for the continuity of identity with continuity of life processes, it seems entirely unmotivated to assume that mental processes are not life processes or to restrict one's attention to processes that minister to the inner organization of the creature, rather than its outward behavior, which may be equally important to its continued survival (e.g. seeking food). Olson (1997) suggests that if we were ever early-stage fetuses, then mental features cannot be part of the self, since we share no mental features with early-stage fetuses. However, even assuming that we were ever early-stage fetuses, this shows only that the self cannot comprise *solely* mental features.

In sum, in my view, the objection from more traditional accounts of the self fails because those accounts also do not (generally) save the identity of self and mind, and moreover, run into problems both in their application and in their supporting arguments because they rely on an unwarranted assumption that the self either solely contains or entirely excludes mental features.

4. Self without mind?

4.1. Taylor and the internalized self

There is another way to resist my argument in Section 2. Rather than reject its relevance by rejecting the relational account of the self, one

could instead object more directly. One might claim that although the distinctness of self and mind follows from 'off the shelf' accounts of the extended, embedded, embodied mind and the relational, embodied self, my response is wrong. Rather than embrace the distinctness of self and mind, so the objection goes, we should *fix* these accounts in order to save the identification of the two.

For example, on the basis that it is in part the emphasis on coupling in the relevant accounts of mind which keeps mind and self apart, one might claim that this emphasis is misplaced, and unduly restrictive, and that the account should be liberalized to allow for the inclusion of those kinds of relations and traits with significant influence on the person. Unless some independent reason is given for so liberalizing the accounts of mind, however, this presupposes that these accounts of mind ought to aim to keep mind as a plausible self-concept.

What reason might be offered for retaining mind as a self-concept? One can, I think, eke out such a reason from the work of Taylor (especially 1989). Taylor traces the development of the modern Western concept of self as a search for the good, and an articulation of what he calls 'moral sources', that is, ultimate ends that are seen as good for their own sakes and against which other ends can be judged.

Much of Taylor's (1989) study traces the 'internalization' of these moral sources – the development of the view that ultimate ends have their basis within us in some strong sense. The idea of internalization is best captured by considering views of rationality, and particularly rational self-rule. Taylor argues that according to the dominant philosophical view of the Ancient Greeks, there is a rational order manifest in the world, and rational self-rule consists in opening oneself up to, and tuning into, this rational order, and one's conduct being governed by one's understanding of this rational order. According to Taylor, by the time of Descartes, the dominant view was that there is no rational order out in the world, but instead, rationality was a feature of our inner processes and procedures for manipulating representations of the world according to evidence and rational self-control consisted in the proper use of evidence and instrumental rationality (working out the best means to achieve one's ends) aiming towards self-sufficiency in conduct and in knowledge (see also Rorty 1979). In as much as the outside world, or 'nature' came to serve once more as a moral source in the later, Romantic era, it was (according to Taylor) by speaking to people from within – an inner voice that guides conduct and is particularly associated with sentiment.

With this process of internalization, according to Taylor, there have developed views of people, and of their place in the natural and social worlds. In his 1989 study, and elsewhere (e.g. 1985, 1999) he critiques many of these views harshly. The process of internalization is associated with the impoverishment of our understanding of our place in the world. Our social relations, joint action and joint cognitive activity, are all mediated by representations in the mind, in a model of consciousness Taylor calls 'monological' (1999) – a model according to which all epistemological contact with the outside world is through inner representations (see also Rorty 1979). It is also, and relatedly, associated with an 'atomistic' view of individuals. According to atomistic views, individuals are, in several sense, prior to the societies of which they are a part – societies are viewed as formed of individuals by the assent of those individuals or as tools that serve the goals of those individuals. The process is also at the heart of the disenchantment of nature. In a narrow sense, the disenchantment of nature is the denial of magic and the supernatural. However, in a broader sense, it does not stop there. The color of the world (including its value, and quite literally its color) is taken no longer to be a property that inheres in the world, on such a view. Mind serves as a container for these now-mere 'secondary' qualities, that are at best projected onto a world that fundamentally consists in bits of dead matter bumping into each other (Taylor 1989; see also Merchant 1980).

Even so, methodologically, his approach is to try to better articulate the underlying ideas and ideals, and save what it is possible to save – or perhaps recognize what it is necessary to retain (see especially Taylor 1991). For example, he seems keen to save the ideals of dignified rational self-control, of autonomy and self-sufficiency, of authenticity and of inner resonance with and commitment to values or ideals. This can be fashioned into a defense of the mind as the self. Taylor holds that the idea of the mind is lacking from Homer, but present in Plato. He holds that the idea of the mind is basically the idea of a unique, unified locus of thought and feeling. Furthermore, he holds that the idea of the mind – of a unified locus – is a prerequisite for the process of internalization, a process which he defends as offering rich discursive moral resources. Even more strongly, in places, he appears to indicate that the internalized view of self *just is* the mind – for example, he suggests that 'mind' is 'the old terminology' for the "'inner" space' often identified with the subject (1999, 33).⁶ If retaining a view of the self as the mind is necessary in order to retain these values and ideals, then we ought at least to try to retain the view of the self as the mind.

It is important to stress that the ideas of inwardness, of an internalized self, and of the mind as the locus of thought and feeling go significantly beyond the recognition of certain basic, obvious truths about human beings. As Taylor says, '[we] can probably be confident that on one level human beings of all times and places have shared a very similar sense of "me" and "mine"' (1989, 112). In particular, the internalized self does not boil down to having a sense of oneself as '[a] single [agent] among others', nor the awareness that one is the person 'whose fate is being decided' by some impending danger, nor an 'understanding of the decisions [one is] called upon to take' (Taylor 1989, 119). Similarly, Taylor claims, there is a sense in which most cultures make a distinction between the inner and the outer – many cultures distinguish between those thoughts, desires and intentions that we keep private and 'inside' and those that we put 'out' into the public domain.

The peculiarity of the modern Western view, according to Taylor, is that idea that we have 'inner depths the way we have hearts or livers, as a matter of hard, interpretation-free fact' (1989, 112). This idea starts with Plato positing an inner locus of thought and feeling, in Taylor's story. The Platonic view is novel in attributing decision-making, individuality, prospective self-interest, thought and feeling to an inner locus, the mind.

Additionally, in the pre-Platonic scheme (specifically the Homeric scheme), external influence essential to an agent's performing an act does not generally diminish 'the merit or demerit attached to the agent' in light of that act (Taylor 1989, 118; see also Teffetteller 2003); instead, external influence is seen as something that can be inseparable from an individual's greatness or blameworthiness. Conversely, in the Platonic scheme, a state of rational self-reflection, capable of surveying all other states, comes to be seen as a location that one can occupy. Occupying this location is seen as the primary good, and special states that are not surveyed and controlled by this rational self-reflective core (including those resulting from strong external influence) are thereby seen as inherent failures (Taylor 1989, 119–120). The shift, then, is to the idea that people have a private inner core, an arena of thought and feeling, that takes responsibility for the basic facts of agency and individuality, and that is (or at least ought to be) capable of rational self-reflection, largely self-sufficient and generally closed to external influence.

⁶He also treats it as evidence that authors have an internalized view of the self when they treat 'mind' and 'self' as interchangeable (1989, 256).

4.2. *Autonomy and rationality as relational*

This argument fails if it is possible to save the ideals and values that we might wish to retain without offering an account of the self that is internalized in such a way that it might deserve the name ‘mind’. This would not be directly in conflict with Taylor’s overarching claims – that internalization affords and afforded a richer moral language, and that the notion of mind may have been necessary for the process of internalization at that time and in that context. It does, at least ostensibly, conflict with a further claim of Taylor’s – that without the idea of a mind, a unique, unified locus of thought and feeling, it is *impossible* to give an account of self-collected reason (1989, 120). He appears to claim that for rationality to serve as an ideal, it must be the case that to experience oneself as a plurality is to experience error and imperfection. I believe that this argument fails because it is possible not only to save, but even to give a better account of, these ideals and values without an internalized self deserving the name ‘mind’.

There are several relational accounts of autonomy, rationality and rational self-control. Indeed, it is the apparent failures of non-relational accounts of autonomy that have motivated many to adopt a relational view of people and selves (e.g. Herring and Wall 2015). Non-relational views of autonomy emphasize freedom from constraint and interference, self-sufficient rational selection of one’s own goals or concordance between one’s first- and second-order desires (both conceived as internal states) as the mark of autonomy. Conversely, relational views emphasize that autonomy is an achievement that one’s relations to others may help to reach and to maintain (e.g. Herring and Wall 2015; Nedelsky 2011; Taylor 1985; Wallace 2019a, 2019b). This is not to deny that one’s relations to others may also hinder autonomy – instead, on this conception of autonomy, part of the challenge is to distinguish between autonomy-facilitating and autonomy-hindering kinds of relations.

Similarly, there are relational views on the offer of rationality and rational self-control. Perhaps the most influential is that of Mead. In Mead’s account, a person’s self is to a large extent a conversation between the outside influences of the social world. In particular, in Mead’s view, in the normal course of development one takes on views from various significant people and groups in one’s life, and these are put into conversation with one another and with one’s spontaneous reactions to these attitudes (Abbott 2020; Mead 1934; Wallace 2019a, 2019b). Through this process, one comes to be able to evaluate and assess one’s

own attitudes, making them fully ‘one’s own’. Such a view has several benefits, one of which is its ability to offer a notion of self that does not fall down to postmodern critiques arguing for the decentering of the subject, often by arguing that selves are essentially sites of political and social conflict. Its notion of self does not fall to such critiques because it *starts* from such an understanding, and *explains* what the self is in light of such a view.

Partly building on Mead, Wallace (2019a see especially 122–125) elaborates a novel, relational account of reflexivity, including rational self-control. Her account understands a person as a ‘community of perspectives’. She highlights that generally, such an understanding of the person has only been used in the explanation of failures of good function – for example, *akrasia* (weakness of the will), self-deception and irrationality. She offers an account of successful function in these terms, opposing any view which posits a central ‘I’ – an executive decision-maker or privileged location from which all can be surveyed.

Her account starts from the notion of an ‘I-position’ or ‘I-perspective’, for example, I-as-son, I-as-dog-lover, I-as-philosopher, and so on. At any given time, several such perspectives are relevant. For example, my Mum just got a new dog, and I-as-son-and-dog-lover went to meet the dog and in doing so see my mother. I-as-philosopher, however, felt I should get on with work. I-as-son-and-dog-lover-and-philosopher was conflicted, but decided nevertheless to go. When my mum and I went for a walk with the dog, not only was I functioning as dog lover, son and philosopher but also as walker – which is to highlight that these functions and perspectives are not solely mental or social.

Reflexivity in general, including rational self-control, is modeled on Wallace’s account as some perspective or combination of perspectives taking a critical eye towards the overall function of the person – that is, of the community of perspectives. There is no central monitor, on such a view, but any I-perspective or combination of I-perspectives may temporarily *act as* an approximation of a central monitor. In sum, according to Wallace, the person ‘is always differentiated and perspectively located’; one is always ‘functioning in a particular way’, and this way should be modeled as a conversation between a community of different perspectives (2019a, 123).

Neither Mead nor Wallace’s account is particularly fatalistic. Growing up the child of a vicar, for example, does not entail that *being the child of a vicar* is a part of your self. Being from the countryside does not entail that you are *essentially* from the countryside. Indeed, while these

may be a source of opinions and perspectives, one is not forced to adopt these opinions or perspectives at all, let alone to such an extent that these features count as part of your self. Identifying self and mind may seem liberating, in that one's social role, background and circumstances seem guaranteed not to be part of one's self. However, this is ensured only by fiat. Mead and Wallace's accounts, conversely, face the issue head on. In doing so, they help to make clear how one can avoid being *essentially* the child of a vicar, or from the countryside – by adopting other opinions and perspectives from other sources, which one can then use as a vantage point for critiquing and liberating oneself from the contingencies of one's background.

4.3. Embodiment

Mead's account, and many relational accounts of autonomy emphasize social relations. This raises the issue of the body, and of how to offer an account of the self, which is embodied as well as relational. There are good reasons, I think, to suspect that the body can be accounted for in similar terms. If the body can also be conceived as a source of opinions and attitudes, as well as a potential helper or hinderer of autonomy, then these views can be generalized to offer a relational, embodied account of the self, and of the ideals Taylor praises.

I suspect strongly that the body can be so-conceived. There is a nice example of such a view in the pre-Platonic language of Homer. Homer lacked any term that can be translated as 'mind', at least consistently (Simon 1978; Snell 1953; Sorabji 1993). On Homer's view, decision making 'is typically a debate within the hero, or between parts of himself' (Simon 1978, 65). Inner life is described as 'a kind of personified exchange between a man and his parts – "The spirit [*thumos*] inside my breast drives me," or "my legs and arms are willing"' (63). This kind of intentionally characterized influence of a person's parts on the person is exactly what is needed to integrate the body's influence on the person with that of the social environment in an account of self and person.

Importantly, and apparently in disagreement with Taylor, experiencing oneself as a plurality is not a *problem* on these views of autonomy and rationality. Autonomy, self-collected rationality and rational self-control are not a matter of ensuring the overall control of a unified inner core. Instead, they can be conceived as consisting in balance between one's many parts and influences, and accordance with the overall pattern of

the person over time (Wallace 2019a, 2019b). To further push the analogy with Homer, external influence does not preclude autonomy or responsibility in Homer's scheme, except when it comes to completely overwhelm all other sources of influence, and override the normal patterns of the person (Teffeteller 2003).

There is another potential role for the body and social world, interestingly indicated in Taylor (1999). There, Taylor critiques 'monological' representationalist models of consciousness, according to which understanding consists in accurate representation of the world, and according to which joint activity is merely co-ordination between individuals via representations. His primary critique (arguably in tension with his 1989; Abbott 2020) is that this vastly overestimates how much of our understanding is articulate or articulable. In particular, he claims that much of our understanding is manifest in practical activity, and that this involves an inarticulable backdrop of habits, embodiment and unarticulated non-representational means of social coordination (he focusses on rhythmic activities). The body and social world, on such a view, shape and form part of this inarticulable backdrop on which our understanding relies, and in which it largely consists. Representational and explicit knowledge, on Taylor's view, are rare islands in a vast sea of understanding.

This insight is important, I think, for Wallace's model of the person as a community of perspectives. One might object to Wallace's view that it does not do away with an executive controller in overall control of integrating and achieving different goals and desires. Instead, one might object, it simply indexes actions, goals, desires and so on to different 'perspectives', and talks of its action in integrating these as 'communication' between perspectives. This objection fails. The reason that it fails is that when I am acting as a son, a philosopher or a walker, the content of these perspectives and the functions they entail, are by no means exhausted by what I can represent. Instead, much of the content and many of these functions are of the inarticulable sort identified by Taylor – embodied, habitual, socially mediated and socially distributed (see also Abbott 2020; Clowes and Gärtner 2020).

4.4. Further issues with self and mind

As such, even if as Taylor claims, internalization was historically necessary in attaining our current discursive moral resources, and views of 'the good', it does not mean that to retain these views we must retain an internalized account of the self on which it can plausibly be identified with the

mind. There are also several other problems with bringing the notion of mind into these discussions. There are good reasons to be suspicious of how helpful the concept mind and category of mental can be to discussions of the self. For example, as discussed above, the exclusivity assumption core to the debate between psychological continuity and animalist accounts relies on the notion (for a more general discussion intimating the unhelpfulness of mind in discussions of self, see Nedelsky 2011, ch. 7).

A more general, and more significant, potential problem is the idea that the view of self as mind and the related view of people as self-transparent, rational, autonomous subjects, was used to justify the long history of racism, colonial oppression and, more generally, moral outrages inflicted on nonwhite races (e.g. da Silva 2007; Lolordo 2019). I worry that it might be illicit to mobilize this issue against *mind* here, because I suspect that whatever self-conception had been adopted by (white, Christian) Europeans at that time, that self-conception would have *somehow* been used to justify the aforementioned moral outrages. In other words, I suspect that in the explanatory order of things, colonization and oppression come first and the wish to justify that according to such-and-such, a self-conception comes second. Where *mind* might nevertheless be criticized as self-concept is in being sufficiently prescriptive about the nature of persons and their selves that it *could* be mobilized in this justificatory role.

There is another, yet more general, reason to suspect that it is a bad idea to try to keep a close relationship between *mind* and *self*, especially if we wish to offer a relational account of the self. The positing of our modern concept of *mind* as a self-concept represented a movement *away* from a Classical understanding of persons that was framed largely in terms of social roles and relations (Brouwer 2019; Lolordo 2019; Van Dyke 2019). This movement was motivated largely by a concern to justify and explain humans' special moral status, and especially their immortality (e.g. Martin and Barresi 2006). This overarching concern can be broken down into two subordinate concerns: a concern to explain what features of humans survived their deaths; and a concern to explain why *only* humans, and not other animals, survive their deaths (Carpenter 2018; Martin and Barresi 2006; Sorabji 1993). These same concerns are also what eventually led *self* largely to fall out of favor as a scientific concept (Martin and Barresi 2006); this is an interesting phenomenon given that it is *mind* more than *self* that represents the culmination of these concerns – more guilty of the sins that led *self* and *soul* into disrepute.

These concerns about the role played by the concept *mind* in our self-conception speak not only to the identification of mind and self but also to the idea that we *have minds*. In Section 2, I argued that the mind is not part of the person but said that this does not preclude the possibility that people *have* minds. The concerns adduced above, however, suggest that there are problems even with the concept of *having a mind*. Another issue is adumbrated in Taylor (1989, 112, 1999, 33). Combining the view that we *have* minds in the way that we have heads or arms, with the view that the mind is (essentially) the arena of rationality, autonomy, and so on, ignores that it takes work and commitment (and often the help of others) in order to attain these ideals. This introduces ‘a tendency ... to read it less as an ideal than as something which is already established in our constitution’ (1999, 33).

I would suggest that we would do better without making use of the concept of *having a mind*. Consider, for example, animal ethics. One might expect that in animal ethics, the concept of *having a mind* might be used to draw lines between systems warranting different levels and kinds of moral consideration, as it perhaps is in ‘folk’ discourse. Instead, however, one primarily finds talk of more fine-grained concepts like agency (the capacity to make choices), moral agency (the capacity to take responsibility for those choices), patienthood (the capacity to suffer), consciousness and rationality. I suspect that this is because the concept of *having a mind* runs together these importantly distinct concepts, which come apart in important ways (see also Schwitzgebel 2020). It is better, I think, where it is necessary to draw such lines, to refer directly to the more specific relevant concept – better, for example, to talk of *being an agent* than of *having a mind*. In light of all these considerations, and absent a strong reason not to do so, I think that it is better simply to jettison talk of mind from talk of the nature of persons and their selves.

4.5. The brain as the internal self?

There is one final objection that I would like to consider. This objection picks up where the last objection started – with the idea that we ought to retain a view of the self and mind such that the two can be identified. However, instead of drawing on Taylor’s moral theory, it instead draws on (perceived) scientific consensus. According to such an account, we ought to retain the theory of the self as mind in order to account for the brain’s significance to the human being. On this view, the body, social relations

and so on, are only important to the person if and in as much as they make an impact on the brain. Here, the brain is identified with the mind, or at least seen as the organ of the mind – that is, the organ characteristically responsible for mental activity, in the way the heart is characteristically responsible for pumping blood around the body or the lungs are characteristically responsible for infusing blood with oxygen.⁷

How, then, can one account for the significance of the brain on an embodied, relational account of the self? One way, of course, is to retain the view that the brain is the mind, to admit that mental properties are extremely significant to the person, and nevertheless to insist that there are other, non-mental properties that are part of the self. In light of the considerations adduced in above, I am not inclined to take this line.

Instead, I think it is better to conceive of the brain as that which enables the various parts of the self – social relations, parts of the body, life processes, emotions and so on – to be put into conversation. The brain, on such a view, helps to orchestrate, integrate and balance diverse influences on the person. This view accords well with recent cognitive science and biology, much of which suggests that the best ‘scientific’ view of a person is as a massive, messy bundle of interdependent and overlapping intentional systems, varying along with several scales of size, complexity and influence.

Such a view also does better justice to the brain’s range of functions. For example, the brain is heavily involved in many non-mental systems for internal regulation and co-ordination, systems whose focus remains squarely on the ‘body’ (e.g. Auletta 2011; Godfrey-Smith 2016; see also Reisman 2016). Earlier, I mentioned that Taylor sees explicit, representational knowledge as rarities within the vast realm of our understanding. A similar view has been expressed by many working in cognitive science, and finds a particularly elegant formulation in Pollack (1994). There, Pollack draws an analogy between flight and cognition. He highlights that flapping, as the most striking feature of biological flight, served as a distraction from the primary purpose of wings and the primary mechanism of flight – a control system for the maintenance of dynamic equilibrium. He claims that ‘[f]lapping is the last piece, the propulsive engine, but in all its furiousness, it blocks our perception that the bird first evolved the aileron principle’. ‘Just like flapping’, he claims, ‘in all its furiousness [symbolic thought] obscures our perception of cognition as an

⁷An account of the brain is of broader significance to the relational account of persons and selves, and especially to my own position. As I indicated scepticism of the idea that we *have* minds; but of course, we have brains, and if brains are minds, then we have minds.

exquisite control system ... governing a very complicated real-time physical system' (114).⁸ Additionally, these non-mental systems play a significant part in steering the person through the world, with a surprising amount of direct influence on our mood and our social behavior (e.g. D'Acquisto 2017; Eisenberger et al. 2010; Moieni et al. 2015; Sapolsky 2017).

The brain is special, on such a view, not because it is the locus of thought and feeling, nor because it is in charge of life processes, but instead because it is a key mechanism by which the interaction between influences from body and world can become a conversation, rather than merely a competition between influences of different strengths. The brain is also the host of (most of) our memory, whose centrality to personal identity can, in turn, be reconceived as stemming from the fact that memory is the primary means by which the influences of one's past can participate in the conversation.⁹

Identifying this conversation with 'mind' seems only to confuse matters. Some of the properties associated with the mind are missing; for example, nothing guarantees any kind of privileged epistemic access to the majority of the conversation, and no central sphere of representations is posited in this picture. Other such properties are distributed much more widely through the person; for example, intentionality and intentional states are granted to the most concrete participants in the conversation such as body-parts.

5. The self is not the mind

The central claim of this paper, made in Section 2, is that the relational, embodied self and extended, embedded, extended mind cannot save the theory according to which the mind is the self – these accounts of self are interested in features with long-term significance to the person, whereas the accounts of mind are interested in entities that enter into dense, reciprocal causal relations with the brain. Furthermore, if these accounts are true, the self is not part of the mind and the mind is part of neither the self nor the person.

⁸Pollack, in my view mistakenly, appears to wish to assimilate the more bodily cognitive functions of the brain to a model of the brain as a control system, and the more social cognitive functions to a more symbolic model, albeit one he sees as '[arising] directly out of the complex behavior of non-linear dynamical systems' (1994, 116).

⁹All this is compatible with the points made by the supporters of the extended, embedded, embodied mind, since the brain need not be the *sole* mechanism that enables this conversation or orchestrates it so as to enable greater sophistication.

I went on to consider several objections over Sections 3–4. In Section 3, I indicated the reasons that I am skeptical both of the psychological continuity account of personal identity, which does not work for everyone, and of the animalist account of personal identity, which appears to presuppose a strangely dualistic view of organisms and biology. In Section 4, I suggested that the concept of *mind* has been generally unhelpful, for many and diverse reasons, in forming an account of people and their selves, and that we would do better to jettison it entirely.

There are two points I hope the reader takes away from Section 4. The first, more sober point is that the relational, embodied account of the self ought to not try to find a new place for it, or to look for friendly theories of mind that appear to ‘resonate’ with this account of the self. The second, more ambitious, more blunt point is that people ought to believe in a relational, embodied account of the self (notwithstanding the caveats indicated throughout the paper), and ought to reject the mind, at least from our politically- and ethically-salient self-conception(s).

Acknowledgements

With thanks to Helen Gough, Alexander Raubo, Josh Lawes and Andy Clark for invaluable feedback on earlier versions of this piece, as well as to an anonymous reviewer instrumental in guiding it to its current form.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Joe Gough  <http://orcid.org/0000-0003-2591-5744>

References

- Abbott, O. 2020. “The Self as the Locus of Morality: A Comparison between Charles Taylor and George Herbert Mead’s Theories of the Moral Constitution of the Self.” *Journal for the Theory of Social Behaviour* 50 (4): 516–533.
- Auletta, G. 2011. *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford: Oxford University Press.
- Behan, D. P. 1979. “Locke on Persons and Personal Identity.” *Canadian Journal of Philosophy* 9 (1): 53–75. <http://www.jstor.org/stable/40231080>
- Bishop, M. A. 2015. *The Good Life: Unifying the Philosophy and Psychology of Well-being*. Oxford: Oxford University Press.

- Blatti, S. 2014. "Animalism." In *Stanford Encyclopedia of Philosophy*, edited by E. Zalta.
- Boyd, R. 1989. "What Realism Implies and What It Does Not." *Dialectica* 43: 5–29.
- Boyd, R. 1991. "Realism, Anti-foundationalism and the Enthusiasm for Natural Kinds." *Philosophical Studies* 61: 127–148.
- Boyd, R. 1999. "Homeostasis, Species, and Higher Taxa." In *Species: New Interdisciplinary Essays*, edited by R. A. Wilson, 141–185. Cambridge, MA: MIT Press.
- Brouwer, R. 2019. "Funerals, Faces, and Hellenistic Philosophers: On the Origins of the Concept of Person in Rome." In *Persons: A History*, edited by A. Lolordo, 19–45. Oxford: Oxford University Press.
- Carpenter, A. 2018. "Illuminating Community: Animals in Classical Indian Thought." In *Animals: A History*, edited by P. Adamson and G. F. Edwards, 63–90. Oxford: Oxford University Press.
- Christman, J. 2009. *The Politics of Persons: Individual Autonomy and Socio-Historical Selves*. Cambridge: Cambridge University Press.
- Clark, A. 2003. *Natural-born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford: Oxford University Press.
- Clark, A. 2013. *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford: Oxford University Press.
- Clark, A., and D. Chalmers. 1998. "The Extended Mind." *Analysis* 58 (1): 7–19.
- Clowes, R. W., and K. Gärtner. 2020. "The Pre-Reflective Situational Self." *Topoi* 39 (3): 623–637. doi:10.1007/s11245-018-9598-5.
- D'Acquisto, F. 2017. "Affective Immunology: Where Emotions and the Immune Response Converge." *Dialogues in Clinical Neuroscience* 19 (1): 9–19.
- Danziger, K. 2008. *Marking the Mind: A History of Memory*. Cambridge, UK: Cambridge University Press.
- da Silva, D. F. 2007. *Toward a Global Idea of Race*. Minneapolis, MN: University of Minnesota Press.
- DeGrazia, D. 2005. *Human Identity and Bioethics*. Cambridge: Cambridge University Press.
- Drayson, Z., and A. Clark. *Forthcoming*. "Cognitive Disability and Embodied, Extended Minds." In *Oxford Handbook of Philosophy and Disability*, edited by D. Wasserman and A. Cureton. Oxford: Oxford University Press.
- Eisenberger, N. I., E. T. Berkman, T. K. Inagaki, L. T. Rameson, N. M. Mashal, and M. R. Irwin. 2010. "Inflammation-Induced Anhedonia: Endotoxin Reduces Ventral Striatum Responses to Reward." *Biological Psychiatry* 68 (8): 748–754. doi:10.1016/j.biopsych.2010.06.010.
- Geach, P. T. 1969. *God and the Soul*. London: Routledge and Kegan Paul.
- Godfrey-Smith, P. 2016. "Mind, Matter, and Metabolism." *Journal of Philosophy* 113 (10): 481–506.
- Greenwood, J. D. 2004. *The Disappearance of the Social in American Social Psychology*. Cambridge: Cambridge University Press.
- Herring, J., and J. Wall. 2015. "Autonomy, Capacity and Vulnerable Adults: Filling the Gaps in the Mental Capacity Act." *Legal Studies* 35 (4): 698–719. doi:10.1111/lest.12094.
- Jirout Košová, M., R. Kopecký, P. Oulovský, M. Nekvinda, and J. Flegr. 2021. "My Friend's True Self: Children's Concept of Personal Identity." *Philosophical Psychology* 34 (1): 47–75. doi:10.1080/09515089.2020.1860209.

- Levinas, E., and S. Hand. 1990. "Reflections on the Philosophy of Hitlerism." *Critical Inquiry* 17 (1): 63–71.
- Lindemann, H. 2009. "Holding One Another (Well, Wrongly, Clumsily) in a Time of Dementia." *Metaphilosophy* 40 (3-4): 416–424. doi:[10.1111/j.1467-9973.2009.01592.x](https://doi.org/10.1111/j.1467-9973.2009.01592.x).
- Locke, J. 1689. *An Essay Concerning Human Understanding*. Oxford: Oxford University Press.
- Locke, J. 1948. *An Essay Concerning Human Understanding, 1690*. Oxford: Oxford University Press.
- Lolordo, A. 2019. "Persons in Seventeenth- and Eighteenth-Century British Philosophy." In *Persons: A History*, edited by A. Lolordo, 154–181. Oxford: Oxford University Press.
- Mackenzie, C., and N. Stoljar. 2000. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. Oxford: Oxford University Press.
- Mackie, D. 1999. "Personal Identity and Dead People." *Philosophical Studies* 95 (3): 219–242.
- Martin, R., and J. Barresi. 2006. *The Rise and Fall of Soul and Self: An Intellectual History of Personal Identity*. New York: Columbia University Press.
- Mead, G. H. 1934. *Mind, Self and Society*. Chicago, IL: University of Chicago Press.
- Merchant, C. 1980. *The Death of Nature*. London: Wildwood House.
- Mill, J. S. 1843/1974. *The Collected Works of John Stuart Mill, Volume III - A System of Logic Ratiocinative and Inductive Part II*. London: Routledge and Kegan Paul.
- Moieni, M., M. R. Irwin, I. Jevtic, E. C. Breen, and N. I. Eisenberger. 2015. "Inflammation Impairs Social Cognitive Processing: A Randomized Controlled Trial of Endotoxin." *Brain, Behavior, and Immunity* 48: 132–138. doi:[10.1016/j.bbi.2015.03.002](https://doi.org/10.1016/j.bbi.2015.03.002).
- Nedelsky, J. 2011. *Law's Relations: A Relational Theory of Self, Autonomy, and Law*. Oxford: OUP USA.
- Olson, E. T. 1997. "Was I Ever a Fetus?" *Philosophy and Phenomenological Research* 57 (1): 95–110.
- Olson, E. T. 2007. *What Are We?: A Study in Personal Ontology*. Oxford: Oxford University Press.
- Olson, E. T. 2021. "Personal Identity." In *The Stanford Encyclopedia of Philosophy, Spring 2021 Edition*.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: OUP.
- Pollack, J. B. 1994. "Book Review: On Wings of Knowledge: A Review of Allen Newell's Unified Theories of Cognition." In *Artificial Intelligence in Perspective*, edited by D. G. Bobrow.
- Rawls, J. 2005. *A Theory of Justice: Original Edition*. Cambridge, MA: Harvard University Press.
- Reisman, M. 2016. "Expanding the Biological Critique to Address the Influence of the Endocrine System." *The American Journal of Bioethics* 16 (2): 52–54. doi:[10.1080/15265161.2015.1120810](https://doi.org/10.1080/15265161.2015.1120810).
- Rorty, R. 1979. *Philosophy and the Mirror of Nature*. Oxford: Blackwell.
- Rose, G. 1981. *Hegel Contra Sociology*. London: A&C Black.
- Sapolsky, R. M. 2017. *Behave: The Biology of Humans at Our Best and Worst*. New York, NY: Penguin Press.

- Schechtman, M. 2018. *The Constitution of Selves*. New York: Cornell University Press.
- Schwitzgebel, E. 2020. "Against the Mind Package View of Minds: Comments on Carrie Figdor's Pieces of Mind." *Mind & Language* 35 (5): 671–676. doi:10.1111/mila.12288.
- Shoemaker, S. 2011. "On What We Are." In *The Oxford Handbook of the Self*, edited by S. Gallagher, 352–371. Oxford: Oxford University Press.
- Shoemaker, S., and S. Swinburne. 1984. *Personal Identity: Great Debates in Philosophy*. London: Blackwell.
- Simon, B. 1978. *Mind and Madness in Ancient Greece: The Classical Roots of Modern Psychiatry*. Ithaca, NY: Cornell University Press.
- Snell, B. 1953. *The Discovery of the Mind*. Oxford: Blackwell.
- Sorabji, R. 1993. *Animal Minds and Human Morals: The Origins of the Western Debate*. London: Duckworth.
- Strawson, P. F. 1959. *Individuals: An Essay in Descriptive Metaphysics*. London: Routledge.
- Taylor, C. 1985. "Atomism." In *Philosophical Papers*, 187–210. Cambridge: Cambridge University Press.
- Taylor, C. 1989. *Sources of the Self*. Cambridge, MA: Harvard University Press.
- Taylor, C. 1991. *The Malaise of Modernity*. Toronto: Anansi.
- Taylor, C. 1999. "To Follow a Rule." In *Bourdieu: A Critical Reader*, edited by R. Shusterman, 29–44. London: Blackwell.
- Teffeteller, A. 2003. "Homeric Excuses." *The Classical Quarterly* 53 (1): 15–31. <http://www.jstor.org/stable/3556479>.
- Van Dyke, C. 2019. "Medieval Mystics on Persons: What John Lock Didn't Tell You." In *Persons: A History*, edited by A. Lolordo, 123–153. Oxford: Oxford University Press.
- Varela, F. J., E. Thompson, and E. Rosch. 2016. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Vold, K. 2018. "Overcoming Deadlock: Scientific and Ethical Reasons to Accept the Extended Mind Thesis." *Philosophy and Society* 29 (4): 489–504.
- Wallace, K. 2019a. *The Network Self: Relation, Process, and Personal Identity*. London: Taylor & Francis.
- Wallace, K. 2019b. "Toward a Reconstruction of Self." *Humana Mente* 12 (36): 189–220.
- Weberman, D. 1999. "Cambridge Changes Revisited: Why Certain Relational Changes are Indispensable." *Dialectica* 53 (2): 139–149.
- Westlund, A. C. 2009. "Rethinking Relational Autonomy." *Hypatia* 24 (4): 26–49. <http://www.jstor.org/stable/20618179>.
- Wilkes, K. V. 1993. *Real People: Personal Identity Without Thought Experiments*. London: Clarendon Press.
- Wilson, R. A., and L. Foglia. 2017. "Embodied Cognition." In *The Stanford Encyclopedia of Philosophy*. Spring 2017 Edition.
- Zahavi, D. 2008. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge, MA: MIT Press.