Phenomenal transparency and the boundary of cognition

Abstract

We argue that phenomenal transparency is necessary for cognitive extension. While a popular claim among early supporters of the extended mind hypothesis, it has recently come under attack. A new consensus seems to be emerging, with various authors arguing that transparency characterises neither internal nor extended cognitive processes. We take this criticism as an opportunity to refine the concept of transparency relevant for cognitive extension. In particular, we highlight that transparency concerns an agent's *employment* of a resource – an agent who merely doesn't consciously apprehend (or attend to) some object may still transparently employ it and, therefore, extend to it. This opens up the possibility for a single object to be transparent *and* opaque for the agent at a given moment in time. Once we understand transparency in this nuanced way, various counterarguments lose their bite, and it becomes clear why transparency is, after all, necessary for cognitive extension.

Keywords: phenomenal transparency, extended cognition, extended mind, cognitive resources

Introduction

Early proponents of the extended cognition thesis argued that *phenomenal transparency*[1] is necessary for cognitive extension (Clark, 2004, 2008; Thompson & Stapleton, 2009; Wheeler, 2005, 2019). More recently, this view has fallen out of favour and a host of authors have advanced arguments against it (Andrada, 2020, 2021; Facchin, 2022; Farina & Lavazza, 2022). These authors claim that both internal and extended cognitive processes can sometimes lack transparency. In this paper, we resist this shift

[1] Phenomenal transparency is distinct from other types of transparency discussed in the literature, in particular procedural transparency and informational transparency. In addition, a concept of transparency is also employed in the literature on AI systems, where it is used to talk about the degree to which the inner workings of such systems are accessible (Andrada et al., 2022). In this paper, we focus squarely on *phenomenal* transparency, which is also sometimes – including by us in this paper – simply called transparency.

away from transparency and argue that, when understood correctly, transparency is necessary for extended cognition.

Generally put, the extended cognition thesis affirms that in certain cases, cognitive processes may involve resources external to the body (Clark & Chalmers, 1998). For instance, when an agent uses their phone to habitually store and retrieve information, her memory may be extended to this device. In this paper, we assume that cognitive extension is possible and discuss one of the conditions required for it.

For an external resource to support an extended cognitive process, it needs to fulfil certain conditions. Among these has long figured the requirement of *phenomenal transparency*, a concept borrowed from phenomenology (Heidegger, 1990; Merleau-Ponty, 2002). Phenomenal transparency – or simply *transparency* – describes how agents, when effortlessly and fluidly employing a resource, tend not to focus on the said resource, but rather on the task at hand. When this happens, the agent no longer consciously apprehends the resource.[2]

The classical example is that of a skilled carpenter (Heidegger, 1990): when employing her hammer, the carpenter doesn't focus her attention on the hammer, but rather on the task at hand. The hammer becomes 'transparent' and the agent 'sees through' it to the task she is trying to accomplish. The object at the centre of her consciousness is not the hammer but the table she is building.

Transparent employment of a resource is argued to shift the subject-world boundary (Thompson & Stapleton, 2009; Wheeler, 2019). The expert carpenter isn't conscious of the hammer but of the table that she is building with it. When the carpenter becomes a master, she sees the hammer no longer as a problem to be solved, but rather as a resource with which to solve problems (Clark, 2004, 2008). The object she is focused on is now the table and not the hammer. Since transparency relates to the subject-world boundary in this way, it has become important in discussions around cognitive extension. Only a transparently employed resource can be a part of an agent's cognitive machinery, and therefore cognitive extension is limited by transparency.

Recently, various authors have argued that transparency is absent in several core cases. In particular, they submit, there are cases in which even clearly cognitive body-internal resources aren't subject to transparency. If that is the case, then we shouldn't expect transparency to be required for body-external cognitive resources either. Secondly, authors argue that transparency is missing in some paradigmatic cases of cognitive extension. They submit that we shouldn't reclassify these cases as non-cognitive, but rather give up on transparency as a condition necessary for extension.

[2] To avoid being unnecessarily wordy, we speak of the *conscious apprehension* of a resource (or even just of *being conscious of* a resource) to mean that an agent consciously apprehends the resource *as an object*. This means, roughly speaking, that the agent is conscious of the resource as an independent object with (at least potential) determinate properties.

We think that such arguments go wrong in important ways. They assume that transparency can be characterised by a mere lack of conscious apprehension of (or attentional focus on) a resource. This forgets that transparency, as it is relevant to cognitive extension, is about the *employment* of a resource. What is crucial, or so we argue, isn't whether an agent consciously apprehends a resource but rather whether the resource's employment relies on this conscious apprehension. Once we see this, it also becomes clear that transparency and opacity aren't mutually exclusive. Consequently, a range of purported counterarguments fail.

Since the counterarguments fail, prior considerations regarding transparency's link to the subject-world boundary – and thus to cognitive extension – apply in full force again. We conclude that cognitive extension requires the transparent employment of resources.

This paper is structured as follows: After this introduction, section 2 introduces the concept of phenomenal transparency and existing arguments for and against its importance to cognitive extension. Section 3 defends a nuanced view of transparency. In section 4, we apply the nuanced view to show why transparency is, after all, necessary for cognitive extension.

Transparency and cognitive extension

The concept of transparency has a storied history in the literature on extended cognition. While early on seen as a crucial – even necessary – ingredient for extension, its status has recently been eroding. We believe this pivot is problematic; however, before we explain why, we must first examine the dialectics of the discussion thus far.

Let's start from the start. Heidegger (1990; see also Clark, 2004), who introduced the concept of transparency in its current guise, points out that when engaging in a skilled activity – such as when a master carpenter employs a hammer – an agent will no longer focus their attention on the tool – but rather on the task she is trying to achieve with it. A skilled carpenter is focused not on the hammer, but rather on the sinking of nails into wood and the making of the table.

Merleau-Ponty (2019) gave us the second core example of the literature: a blind person navigating the world with her white cane. Her focus, he says, isn't on the tactile vibrations that the cane elicits in her hand but rather on the world at the tip of it. She will, for instance, encounter the pavement as her cane bumps into its edge.

If the resource isn't at the focus of attention during its transparent employment – what is? In short, when using a resource transparently, we focus on the task for which it is employed. This idea has often been described by saying that when a resource is used transparently, we see *through* it. It is in this sense that the resource becomes transparent. A blind person navigating the world with a cane doesn't focus on the vibrations transmitted to her hand by the cane, but rather on the world they perceive

through it. A skilled carpenter employing a hammer is focusing on joining pieces of wood rather than on how she needs to employ the hammer to do so.

The two examples illustrate an important distinction among cases of transparency. The carpenter employs the hammer to *act* on the world; she is building a table and is using it as a means to this end. When employing the hammer *transparently*, she consciously apprehends pieces of wood, nails, and the table she is making (that is, not the hammer). In contrast, the novice carpenter sees herself confronted with a second task: the proper employment of the hammer. She needs to consider how she would like the wood to be joined and then *infer* the appropriate way of striking the nail. The skilled carpenter doesn't focus on the hammer (and it thus becomes transparent to her) whereas the novice must pay attention to it (and employs it opaquely).

Most of what applies to the carpenter also applies to the blind person – but with an important difference: she employs her cane to *perceive* the world. When we want to learn about the world, we sometimes use body-external resources, and these can be employed transparently or opaquely. When the blind person is using her cane automatically and effortlessly, she perceives that she has encountered the pavement. A novice, in contrast, perceives state changes in the tool she is using (for instance, the changes in the white cane pressing against her hand) and uses that to *infer* the information she is really after (that is, from the vibrations in her cane, she infers that she must have encountered the pavement).

Some of extended cognition's early proponents think that employing a resource transparently is necessary – or a reliable indicator – for cognitive extension. Clark (2004, 2008) juxtaposes transparent and opaque technologies and shows how the transparent use of resources underlies the kinds of cases in which he thinks cognition extends. Similarly, Wheeler (2019) says that when an external resource is transparent, 'a necessary condition [is] met for its constitutive incorporation into the user's mental machinery' (p. 862). Finally, Thompson and Stapleton (2009) propose a *transparency constraint* according to which an external resource must be employed transparently if it is to become part of an agent's cognitive system.

Various authors have linked transparency to the constitution of the subject-world interface, and it's because of this link that they think it is necessary for cognitive extension (Clark, 2004; Thompson & Stapleton, 2009; Wheeler, 2019). The skilled carpenter's attention is focused on the table; that's the object of which she is conscious. Likewise, the blind person perceives the world at the tip of her cane; this is the world from which she conceives herself to be distinct. In contrast, the hammer and the cane are part of the subject (or 'body-as-subject' as Thompson and Stapleton (2009) write) which allows them to perceive and act in these ways.

A recent account by Grush and Springle (2019) helps clarify just how transparency helps constitute the subject-world interface. Let's re-examine what happens when the blind person encounters the pavement: When the tip of her cane bumps against the pavement, vibrations are caused that travel upwards to the user's hand. There, the

signals are picked up by touch receptors, transduced, and sent along nerves travelling to the brain, where further processing is taking place. The agent is thus connected to the world by a *causal chain* that enables her to gain information about it.

Grush and Springle focus on explaining why certain agents perceive objects that are more (the pavement) or less (the white cane) far along the causal chain. They argue that perception depends on internal models, which map from internal causes (say, certain neural signals) to external causes.[3] When an agent lacks experience with a type of causal chain, her internal model may map neural signals only to causes quite close to her. Hence, a novice's internal model might map a given neural signal to the fact that the cane in her hand is vibrating in some specific way. Such a user will then, in a second step, need to infer what worldly cause might be behind those vibrations. Here, the agent attributes properties to an object (the cane) and then uses her knowledge of those properties to make further inferences. In contrast, an experienced cane user may map neural signals to causes further out in the environment. A given neural signal may let her know that she has encountered the pavement. She doesn't need to consciously *infer* that she's encountered the pavement.

Sophisticated internal models map directly to more distant causes, skipping intermediary causal links. Those causal links that are skipped in such a way are transparent to the agent. The agent *perceives* the object and properties to which the internal model maps. The further out in a causal chain the object is, the further out an agent will experience an interface between herself and the objective world. Because intermediary steps are transparent, they aren't consciously apprehended as objects and are rather experienced – if they are experienced at all – as a part of ourselves. A novice cane user is conscious of the cane in her hand whereas the expert cane user is conscious of the pavement at the tip of her cane.

Recently, a number of authors (Andrada, 2020, 2021; Facchin, 2022; Farina & Lavazza, 2022; Smart et al., 2022) have made it clear that they aren't enthused by arguments such as the one just presented. Their arguments generally involve cases which purport to show that transparency cannot be necessary for extension. Two types of cases are generally presented: (1) cases where internal cognitive resources aren't employed transparently and (2) cases where extended cognitive resources aren't employed transparently. We'll look at them in turn.

Andrada (2020) argues that even our own bodies may be opaque to us, and our bodies are generally seen as *the* paradigmatic example of transparently employed resources. She shows how many women, socialised in societies with patriarchal value systems, perceive their own bodies as burdensome. We agree that an object perceived as burdensome isn't transparent: it's at the focus of attention and it figures as an object in consciousness. Facchin (2022) mentions a case originally developed by Dennett

[3] Grush and Springle's (2019) proposal accounts for both perception *and* action. To fully understand the true force of their argument, it is necessary to consider them jointly. However, for the point we are making in this paper, it suffices to look at perception alone.

(1981) in which an agent's brain is transplanted into a vat and hooked up to the rest of the body with a wireless communication system. We are then asked to imagine the agent looking at their own brain. Clearly, that brain still powers the agent's cognition, but just as clearly, the brain is what the agent is attending to and is thus opaque.

Cases such as these are used in conjunction with the parity principle to undermine transparency's necessity for cognitive extension. The parity principle has been widely employed in arguments for extended cognition (see Clark & Chalmers, 1998) and says, roughly speaking, that an external resource should be deemed a part of the cognitive machinery if it functions as a process that would be called cognitive, were it instantiated by an internal resource. Authors such as Andrada (2020, 2021) and Facchin (2022) argue that because internal cognitive resources aren't necessarily transparent, we cannot demand that this should be the case for external resources. Opaquely employed resources may therefore extend our cognition.

The second kind of argument against transparency's necessity for cognitive extension is straightforward: pick a case in which an agent's cognitive processes extend to a body-external resource and show how the agent isn't employing the resource transparently. When playing Tetris, agents will often rotate the falling blocks (*zoids*) not just to orient them such that they will seamlessly slot in at the bottom, but also simply to figure out whether a zoid would fit when rotated in various ways (Kirsh & Maglio, 1994). According to this line of thought, players are replacing body-internal mental rotation with mental rotation processes that reach out into the world. Facchin (2022) agrees that such cases evince a form of *cognitive* processing and then argues that this is the case even though the relevant resource isn't employed transparently. After all, the players are focusing their attention on the zoids; it's these that figure as objects in the players' consciousnesses. If some paradigmatic cases of extended cognition aren't subject to transparency, then clearly transparency isn't necessary for extension.

At this point, we want to highlight two crucial assumptions guiding the above kinds of argument. First, there's a focus on how, when an agent employs a resource transparently, she doesn't consciously apprehend the resource as an object (or, alternatively, she doesn't attend to the object). It's when an agent employs a resource without being intentionally directed at it that it's being employed transparently (see Facchin, 2022). Similarly, Andrada writes that 'consciously encountering an external resource does not, by itself, prevent it from being incorporated into an agent's system' (p. 4706), implying that it's the lack of such a conscious encounter that engenders transparency. Even those who think that transparency is important for cognitive extension at times fall prey to such assumptions. Wheeler (2019), for instance, says that transparent resources aren't cognised as 'identifiable bearers of determinate states and properties' (p. 859). All of these authors seem to assume that transparency is sufficiently characterised by the lack of conscious apprehension of (or attention to) the object.

The second assumption is that transparency and opacity are mutually exclusive.

Note the structure of the above arguments: they begin with cases in which an agent is using some resource opaquely, move from this to the claim that the agent isn't using the resource transparently, and finally conclude that cognitive processes cannot, therefore, extend to the resource. For instance, because the Tetris player attends to the zoids, she cannot be employing them transparently and therefore no extended cognitive processes characterise her interaction with the game.

We show in the next section that there are subtle mistakes in both of these assumptions. Transparency, as relevant for cognitive extension, concerns the *employment* of a resource in a cognitive task. Such employment may be characterised by transparency while the resource in question also figures as an object in consciousness (viz. Wheeler, 2019). This means that it's possible to employ a resource transparently while also encountering it opaquely.

A nuanced view of transparency

In this section, we work out some important nuances that the existing literature on transparency has missed. As we will show in a later section, once these nuances have been understood, the importance of transparency for cognitive extension can be properly appreciated.

We begin by making two preliminary suggestions concerning the transparent use of *cognitive* resources. First, we think the transparent employment of such resources can be elucidated by the distinction between representational vehicles and content. Second, transparently employed cognitive resources do not always neatly fit into either of the two standard categories discussed in the last section; they aren't always about the employment of resources to act on or perceive the world.

Our first suggestion is that when we're using a cognitive resource transparently, our focus is on the *content* of the relevant representation to the exclusion of the representation's vehicle properties. Consider an agent who, using a calculator automatically and effortlessly, wants to find out the sum of two numbers. If she is using the calculator transparently, she may consciously intend to find the sum and then apprehend the resulting sum. The agent doesn't attend to the calculator's vehicle properties (that is, the physical properties with which it enables mathematical operations). The agent doesn't – and doesn't need to – focus on the calculator's buttons, its screen, or its inner workings.

Opaque use of a calculator is different. Here, the user may have the same end goal as before (that is, calculating a sum) but, in order to achieve it, she *additionally* needs to think about changing the states of the calculator (pressing certain buttons, reading the screen, and so forth). Unlike in the transparent case, she needs to attend to and consciously manipulate the physical properties of the calculator to apprehend the contents she is interested in. When employing a cognitive resource opaquely, the end

goal cannot be achieved without first achieving some intermediary task that requires attending to the physical properties of the required tools.

On to the second preliminary point: sometimes an agent transparently employs a cognitive resource but does so neither for action nor perception. This applies, for instance, to the case of the calculator, which – unlike these earlier cases based on action or perception – doesn't seem to be used to interface with the world. When the agent becomes skilled in their use of the calculator, a process that once crossed the subject-world interface becomes a cognitive process whose inputs and outputs are also located within the boundaries of the subject.

This illustrates that with certain cognitive resources, a subject-world interface may disappear *without giving rise to a new one*. A novice user of a calculator encounters an object, the calculator, and manipulates it to achieve her task. However, it's not the case that once skilled, the subject-world interface shifts for her, so that she encounters the world *through* the calculator. No, certain tasks take place entirely within the confines of our cognitive systems. Now, this isn't to say that no shift in the boundary between the subject and world has taken place; it's just to say that this boundary isn't an interface through which the world is encountered.

This can lead to confusion because authors may be looking for a shift in the subject-world interface to evaluate whether a cognitive process has become extended, and here this doesn't happen. We'll get to this in more detail in the next section. For now, it suffices to note that a cognitive process may become extended without the establishment of a new subject-world interface.

With these preliminaries out of the way, we now turn to the core of our proposal for a nuanced understanding of transparency. In particular, we want to respond to authors who assert that we cannot consciously apprehend some object while also employing it transparently. As we will see, it isn't incoherent to combine transparency and opacity about a single resource at a given moment in time.

It's easiest to see this with an example: Imagine a cognitive scientist who is studying the brain. Lacking suitable test subjects, she has hooked up her brain to a scanning device and is observing her evolving brain states on a monitor. Is she transparently or opaquely employing her own brain?

In a sense, her brain states are opaque to her. After all, she is seeing them evolve on a screen in front of her, she is attending to them, and they are very much visible and present at the centre of her conscious experience. Note, however, how her *employment* of her own brain seems to be entirely independent of what she's seeing on the monitor. Nothing of relevance changed in how she employs her own brain when she sat down and started studying the monitor. If she had been employing her brain transparently before she started looking at it, then it seems she must be employing it transparently now, too.

For those who aren't convinced, let's suppose that the monitor shows another human being's brain states. In this case, the cognitive scientist is clearly employing her

own brain transparently. But how has her employment of her own brain changed between the original case and the present one? We don't think there is a relevant difference, and we submit that she's employing her own brain transparently in both cases. If you're *still* hesitant to agree, imagine our scientist looking at a brain she *thinks* belongs to another person, but which is in fact her own. Phenomenologically speaking, nothing might distinguish this case from the one in which she is looking at another person's brain. Clearly, when looking at someone else's brain, the scientist is employing her own brain transparently. And if it's possible that looking at one's own brain and looking at another person's brain give rise to the same phenomenal states, then – given that transparency is a phenomenological notion – it is possible to look at one's own brain without thereby losing transparency.

The kind of transparency relevant to cognitive extension is about the *employment* of a resource, and it's possible to transparently employ a resource while utilising said employment to (opaquely) think about the very same resource. As we've seen above, transparency arises when attention is focused on something other than the vehicle properties of the resource. Note how the cognitive scientist is still automatically and effortlessly employing her brain to 'see through' it. She doesn't manipulate any of its vehicle properties to form the various beliefs she has about her own brain. She gains those beliefs *because of* her effortless and automatic employment of her own brain.

Merely being conscious of (or attending to) the relevant resource therefore doesn't spell the end for transparency. For opacity to exist to the exclusion of concurrent transparency, the agent needs to *employ* the resource with the help of the properties of which she is conscious. Here's an example of how that might happen: Our cognitive scientist could use information about her brain to suppress her own alpha waves. Suppressing these isn't something we can just will to happen, but it has been shown that when we employ neurofeedback devices – such as a monitor on which we can see our evolving brain states – we can, with practice, will those brain states to change in the required manner (Bagherzadeh et al., 2020; see also Clark, 2015). If our cognitive scientist did *that*, she really would be employing her own brain opaquely.

The task on which our scientist is focused when attempting to influence her own alpha waves involves the manipulation of certain physical properties, namely the alpha waves she sees on the monitor. She is attempting to bring about a change in her cognitive states by manipulating an object, namely her brain. This is unlike when she's employing her brain transparently, which doesn't require considering vehicle properties at all and where simply attending to her mental states suffices.

Hence, contra some recent arguments (see Facchin, 2022), it's not incoherent to speak of the transparent employment of a resource in combination with a concurrent opaque conscious apprehension of the resource. As we've shown, we can use a cognitive resource automatically and effortlessly to think *about that very resource*. We may *think with* cognitive resources and we may *think about* cognitive resources, and we must take care to keep the two apart.

You may wonder whether neurofeedback devices aren't a problem for our claims. After all, it seems that when our cognitive scientist employs her brain opaquely, she isn't just conscious of her brain, but rather uses her knowledge of her brain's physical properties to manipulate it. This is, therefore, not the kind of case in which a resource is transparently employed to think about that very resource. If, as we have repeatedly claimed, transparency is required for cognitive extension, then neurofeedback cases would seem to drive us to say that the brain is here expelled from the cognitive system – and that is patently absurd. An appropriate response to this challenge will need to wait for a moment, but – to avoid keeping you on the edge of your seat – we'll indicate here that the resolution depends on the fact that a given resource can be involved in multiple cognitive processes, only some of which may be transparent.

We hope to have clarified the nature of transparency as relevant to the employment of cognitive resources. Transparency is about the *employment* of a resource to achieve some task. Whether transparency characterises our relation to some resource depends on whether attention is paid to the resource to achieve the task. Moreover, and relatedly, it is possible to transparently use a resource to (opaquely) think about that very resource.

Transparency and cognitive extension, revisited

In this final section, we turn to the main question: how, if at all, is phenomenal transparency required for cognitive extension? In contrast to much of the recent literature, we argue that transparency – understood in the nuanced way discussed in the previous section – is a necessary requirement for cognitive extension.

Let's begin with a look at a well-established example from the literature on cognitive extension: tactile-visual sensory substitution (TVSS) devices (see for instance, Clark, 2004). Such devices allow blind people to perceive the locations of objects in space in a similar way to how sighted people perceive the world using their visual systems. A device is attached to some part of the body – for instance the forehead – and hooked up to a camera. The device has an array of (electro- or vibrotactile) activators which are triggered based on the information captured by the video camera. The device converts the camera's signal into a tactile stimulus, which agents can then use to gain spatial information about their environments. For instance, a tactile sensation at the centre of the array might indicate an object located in front of the agent.

When an agent is new to such a device, gaining spatial information from it requires the agent to focus on the tactile sensations and use these to infer information about, say, the distance to a chair they're facing. In cases such as this, the agent is using the resource opaquely; she is focusing on the tactile stimuli and is using her knowledge of the device's functionality (in particular, her knowledge of how certain stimuli correlate with object location) to achieve her cognitive task (that is, inferring the location

of the chair).

With time, agents begin to use TVSS devices effortlessly and automatically. They no longer focus on their tactile experiences to infer the spatial organisation of their surroundings. When they enter a novel room, they no longer attend to the stimuli on their foreheads to infer that there's a chair blocking their path; rather, they simply experience the chair being located in their path. When they wonder about the locations of objects in their surroundings, their focus is on the content of their 'visual' experience rather than the tactile sensations. They no longer consciously experience the tactile sensations or the fact that they're employing a TVSS device – they are using the device transparently and perceive their surroundings 'through' it.

Now, let's say we have established that some agent is using her TVSS device transparently. Does this entail that her cognitive processes extend to that device? Most authors, if not all, do not take transparency to be a sufficient condition for cognitive extension and therefore answer the question in the negative. There exist additional conditions – for instance Clark's (2010) glue and trust conditions – that may fail to obtain even if transparency obtains (Facchin, 2022; Wheeler, 2019). These additional conditions aren't the focus of this paper, and we will not discuss them further here. The literature focuses – and so do we – on whether transparency is necessary for cognitive extension.

Before looking at the claim regarding transparency's necessity for cognitive extension, we need to first specify how exactly we are to understand it. Only then can we assess the implications of the cases that are used to argue against it. As we will see in a moment, we may distinguish between two types of view on cognitive extension, and it turns out that if transparency is necessary, it is necessary in very different ways depending on which view we espouse.

Call an account of cognitive extension *dispositional* when it focuses on the extension of dispositional cognitive states. Clark and Chalmer's (1998) original account serves as a useful prototype. They focus on the extension of dispositional beliefs (such as Otto's belief regarding the location of the MoMA) and argue that such a belief may be (partially) realised in body-external objects (such as Otto's notebook). If we subscribe to such an account, what matters first and foremost isn't how some agent is *actually* employing a resource, but rather how they are disposed to do so. Otto has an extended belief not because he is right now employing the notebook in a certain way, but because he is so disposed. Were he interested in the location of the MoMA, he would pick up his notebook, look up the address, and generally engage in behaviours indicative of the possession of a dispositional belief.

When we apply this thinking to TVSS devices and the importance of transparency, it becomes clear that if transparency is necessary for extension, it's necessary in the sense that the user must be disposed to use the device transparently. If the kind of disposition at play here is a probabilistic disposition (rather than a 'surefire' disposition, see Choi & Fara, 2018), exemplification of the disposition doesn't entail that the agent

never uses the device opaquely. Just as someone may be disposed to cycle to work even if they take the bus on isolated occasions, it may be possible to be disposed to use a device transparently without always doing so. Thus, on dispositional accounts, if transparency is necessary for extension, it is a modal condition: what matters isn't whether the agent actually employs the resource transparently, but whether she is disposed to do so.

We may, in contrast, understand extended cognition as *dynamical*, that is to say, as driven by how agents dynamically recruit resources as they go about their lives, using whatever is most useful to accomplish the tasks at their hands. Wheeler (2019; see also Clark, 2007), for instance, seems to subscribe to such a view. Here, an agent extends as she begins to employ resources in a certain way and then shrinks down again once she stops employing the resource. Agents dynamically expand and contract. According to this view, if transparency is necessary for extension, then an instance of non-transparent resource use entails that the agent fails to extend to said resource.

We do not have a stake in the debate between the two views on cognitive extension (indeed, the two aren't mutually exclusive). We simply want to highlight that even if it is established that transparency is necessary for extension, we can't always infer from an individual case of non-transparency that an agent's cognition doesn't extend. Such inferences are only valid if we subscribe to a dynamical account.

Of course, we still need to answer the question of whether (actual or dispositional) transparency is necessary for cognitive extension. Recall that transparency is generally employed in parity-style arguments: internal mechanisms which are deemed cognitive are said to be transparently employed, and thus (as long as other conditions are also fulfilled) we should also consider external resources as part of our cognitive machinery if they are so employed. However, just because some cognitive process has a certain property, it doesn't follow that *any* cognitive process needs to have this property. Whence the question about necessity: what makes transparency necessary for cognition (be it with body-internal or body-external resources)?

We have already seen the answer to that question: transparency is important for the establishment of the self-world interface. As the name of the concept implies, when a resource is employed transparently, then we 'see through' it. The resource itself becomes part of the machinery with which we perceive or act on the objective world, a part of us *qua* subject. Hence, transparency is relevant for cognition insofar as cognition is what happens within the boundary of the agent (at least on the kinds of view of cognition underlying the extended cognition hypothesis) and only transparently employed resources may be located within that boundary.

To remind ourselves how this works in the case of cognitive extension, let's take another look at TVSS devices. To get the information that there's a chair in front of her, the novice TVSS device user needs to focus on tactile stimuli on her forehead and infer from these the presence of the chair. To do *that*, she needs to attribute certain properties to a certain object (that is, she needs to attribute properties to the TVSS

device). It's in this conscious apprehension of an object (and its properties) that a boundary between subject and world is instituted – the TVSS device is part of the world perceived by the agent rather than part of the machinery with which she perceives it. And that means the TVSS device cannot be a part of the agent's cognitive system.

According to Grush and Springle's (2019) account, the novice's internal model maps neural signals to information about how the TVSS device is touching her forehead. The user may use this information to *infer* object states further out in the causal chain – she may infer that a chair has caused her video camera to pick up some signal, which has then been transformed to cause the tactile sensation she is experiencing. This is unlike the transparent case, in which the internal model maps directly to causal links further out in the causal chain (namely, to objects in the environment). Thus, without transparency, the system attributes properties to one object (the TVSS device) whereas, with transparency, properties are attributed to another (the chair). Whereas in the first case, the TVSS device is experienced as an object (and thus cannot be a part of the agent as a subject), in the second case, the TVSS device is part of the machinery with which the agent can perceive the chair.

As previously mentioned, authors who are opposed to such conclusions have formulated two main kinds of counterargument. These generally begin from counterexamples involving clearly cognitive body-internal processes that are opaque (showing that cognitive processes need not necessarily be transparent) or paradigmatic cases of cognitive extension which are subject to opacity (showing that extension doesn't necessitate transparency).

First, given the above considerations regarding the two general types of account of cognitive extension, individual cases aren't always decisive. If we subscribe to a dispositional account of cognitive extension, extension may be possible even when an agent uses some resource non-transparently at times. However, as we'll show in the remainder of this paper, the problems go deeper and also concern dynamical accounts of cognitive extension.

We have, in fact, already dealt with counterarguments of the first kind. In the previous section, we've looked at a cognitive scientist who uses a neurofeedback device – a case used by Facchin (2022) to argue against transparency's necessity for cognitive extension. We agree that when this scientist employs her device, she is thinking about her brain opaquely. We also agree that it would be absurd to claim that the brain is expelled from the agent's cognitive system when it's used in such a way. However, we disagree with Facchin when he concludes that transparency cannot be required for a resource to become a part of an agent's cognitive system. As we've said previously, the cognitive scientist still employs her own brain transparently – it's just that this transparent employment is targeting the brain itself.

If we employ Grush and Springle's (2019) account, we can see that this is a case of a causal chain that bends back onto itself. Brain activity is picked up by sensors,

processed, displayed on a monitor, received as light by the user's retinas, and finally processed by the brain to realise various beliefs. Plainly, the brain figures twice in the causal chain, and only the first instance is opaque to the agent. The final processing that realises the agent's beliefs is still transparent to her. Thus, this purported counterexample fails as the agent employs her cognitive machinery transparently even here.

This concurrency of transparency and opacity also characterises – as parity-style arguments imply – standard cases of cognitive extension. Imagine a seasoned user of a TVSS device facing a mirror. What she 'sees' in the mirror is the TVSS device strapped to her forehead that enables her to see in the first place. Nothing in her employment of the resource necessarily changes – in fact, her *seeing* the device (rather than inferring its presence based on tactile stimuli) implies that the device is still used transparently. Thus, also in the case of body-external resources, we have cases in which opacity of a resource doesn't preclude concurrent transparency.

The second type of counterargument comes in the form of paradigmatic cases of cognitive extension in which an agent employs the relevant resource opaquely. This is what is argued to happen in the case of Tetris, where players use extended cognitive processes to mentally rotate the falling zoids while clearly, at the same time, attending to these falling blocks. The case seems to show how an agent can be conscious of some resource while her cognitive processes extend to that very same resource.

Answering this second challenge requires a little more work than the first one. The difficulty lies in the fact that the zoids are employed in an opaque way to achieve the task of slotting them into place. Thus, unlike in the above case of the neurofeedback device used to form beliefs about one's own brain, the zoids don't just figure as objects in consciousness, but the player is using her knowledge of the zoids to employ her resource.

In the present case, a resource (a computing device running Tetris) is supporting multiple processes, only some of which are characterised by transparency. Note, first, an interesting lack in the description of the phenomenology of the case: the relevant study (Kirsh & Maglio, 1994) looked at players with at least an intermediate skill level, and these are likely able to employ the computing device transparently (that is, they do not need to attend to the buttons to operate the device). They are faced with the task of efficiently placing zoids, without having to additionally solve the task of making the zoids move as they intend them to move. To place the zoids efficiently, players need to know how the zoids look in various orientations so that they can find complementary gaps at the bottom of the screen. Players employ the computing device to do two things: first, they must consider how zoids fit at the bottom of the screen given various orientations, and second, they need to slot the zoids into place. Kirsh and Maglio call the first an epistemic and the second a pragmatic action (Kirsh & Maglio, 1994).

We submit that to achieve her first task, a player transparently employs the device to rotate the zoids, which doesn't require consciously apprehending the zoids and

their properties. Unfortunately, only further empirical research can provide evidence to support this claim, but we think that the players' apparent unawareness regarding the goals of their epistemic actions makes it at least plausible. The second task – slotting the zoids into place – does seem to require consciously attending to the zoids. Thus, the confusion surrounding this case may be due to the fact that a single external resource is involved in multiple processes – and is only employed transparently in some of them.

We have now come across a second way in which opacity and transparency aren't mutually exclusive. Already in the previous section, we noted that in some cases, a cognitive resource may be transparently employed to consciously apprehend that very same resource. We argued that such cases may involve extended cognition as long as the resource is still *employed* transparently. The present case is different: here, a resource *is* employed opaquely: when slotting a zoid into place, the agent uses her knowledge of the shapes on the screen to slot them into place. However, the agent is also employing the device transparently: when it comes to gaining information about how the zoids look in various orientations, the computing device is used transparently. Here, opacity and transparency co-exist because they characterise two different processes that make use of the same resource.

Analysing the case in such a way has intriguing implications. First, it shows that to properly make sense of transparency and its importance to cognitive extension, we need to subscribe to a process-based account of cognitive extension (see Menary, 2012). Such an account doesn't focus on whether certain objects are a part of the cognitive machinery, but rather whether certain processes reaching through those objects can be properly called cognitive. Second, when we subscribe to such a process-based view, it becomes possible that certain resources are involved in a variety of processes, only some of which might be cognitive. It's no longer given that we can give an absolute answer to the question regarding some resource's inclusion in the cognitive machinery (or, alternatively, it no longer follows from a resource's inclusion in the cognitive machinery that *all* processes running through it are cognitive).

Such an account also allows us to iron out a wrinkle in the last section. There, we noted that if our cognitive scientist were to use the information she gains through her neurofeedback device to manipulate her brain's alpha waves, she would be employing her brain opaquely. Given our argument, she seems to thereby chuck her brain from the cognitive system – and that is absurd.

We are now able to account for this case. Various cognitive processes run through the cognitive scientist's brain, and most of these are unaffected when she starts to (opaquely) manipulate her brain states. What changes is that there's now an additional process that takes the cognitive scientist's own brain as an object, and *that process* employs her brain opaquely. The brain is still a part of the agent's cognitive machinery insofar as many of the agent's cognitive processes run through it and most (but not all!) of the processes running through it are the agent's cognitive processes.

The brain still underlies much of the agent's cognition and, for the most part, she still uses it transparently.

You might object that such a process-based account still falls short. Facchin (2022) argues that it's not clear how a new subject-world interface comes to be when playing Tetris. The first process (mentally rotating zoids) doesn't seem to be such that a new interface appears. And the second process (slotting zoids into place) is such that the zoids are firmly on the object side of the subject-world boundary.

We think that we may understand what's happening once we see an analogy between the first process (rotating the zoids) and the previously discussed case of the calculator. Recall that the skilled user of a calculator may extend their cognitive processing to it without thereby using the calculator to act on or perceive the world. Similary, the skilled Tetris player internalises the process of rotating zoids (making it a case of *mental* rotation) without thereby employing the zoids (or the computing device) to perceive or act on the world. Rather than being a case in which a subject-world interface is shifted, it's a case in which the interface disappears.

The case is complicated by the fact that there exists – as we've seen above – a second process which *does* employ the resource opaquely. Because of that process, a subject-world interface persists such that the zoids are consciously apprehended as objects. Thus, there is a first (extended cognitive) process that doesn't give rise to a new subject-world interface, and there is a second process that institutes a subject-world interface.

Things have gotten complex – but such is cognition. What is important for our purposes is that no matter the complexity, one constant remains: for a cognitive process to extend to some resource, the agent must employ the resource transparently. It doesn't follow that this resource cannot be an object of cognition (as when seeing one's TVSS in the mirror) or that there cannot be other processes which employ the object opaquely (as when playing Tetris). Transparency is required – and that's it.

Conclusion

Transparency is necessary for cognitive extension. However, as recent counterarguments have shown, the case isn't as straightforward as previously thought.

We have shown that the kind of transparency relevant to cognitive extension is about the *employment* of resources. For cognitive extension, an agent mustn't apply her knowledge of the resource to manipulate it. Transparency needs therefore not be absent just because an agent consciously apprehends (or attends to) a resource. An object may (opaquely) figure as an object in our conscious experience while at the same time being used transparently. Moreover, a resource may help instantiate multiple processes, and only some of these may involve transparency. There are therefore at least two ways in which an agent may use a resource transparently and opaquely at

the same time.

Bibliography

Andrada, Gloria. (2020). Transparency and the phenomenology of extended cognition. *Límite: Revista de Filosofía y Psicología*, *15*.

Andrada, Gloria. (2021). Mind the notebook. *Synthese*, *198*, 4689–4708. http://dx.doi.org/10.1007/s11229-019-02365-9

Andrada, Gloria, Clowes, Robert W., & Smart, Paul R. (2022). Varieties of transparency: Exploring agency within AI systems. *AI & SOCIETY*. http://dx.doi.org/10.1007/s00146-021-01326-6

Bagherzadeh, Yasaman, Baldauf, Daniel, Pantazis, Dimitrios, & Desimone, Robert. (2020). Alpha synchrony and the neurofeedback control of spatial attention. *Neuron*, *105*, 577–587.e5. http://dx.doi.org/10.1016/j.neuron.2019.11.001

Choi, Sungho, & Fara, Michael. (2018). Dispositions. In Edward N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Fall 2018). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2018/entries/dispositions/

Clark, Andy. (2004). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. Oxford University Press.

Clark, Andy. (2007). Soft selves and ecological control. In *Distributed cognition and the will: Individual volition and social context* (pp. 101–122). MIT Press.

Clark, Andy. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

Clark, Andy. (2010). Coupling, constitution, and the cognitive kind: A reply to adams and aizawa. In Richard Menary (Ed.), *The extended mind*. MIT Press.

Clark, Andy. (2015). What 'extended me' knows. *Synthese*, *192*, 3757–3775. http://dx.doi.org/10.1007/s11229-015-0719-z

Clark, Andy, & Chalmers, David. (1998). The extended mind. *Analysis*, *58*(1), 7–19.

Dennett, Daniel. (1981). *Brainstorms: Philosophical essays on mind and psychology*. Bradford Books. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=48930

Facchin, Marco. (2022). Phenomenal transparency, cognitive extension, and predictive processing. *Phenomenology and the Cognitive Sciences*. http://dx.doi.org/10.1007/s11097-022-09831-9

Farina, Mirko, & Lavazza, Andrea. (2022). Incorporation, transparency and cognitive extension: Why the distinction between embedded and extended might be more important to ethics than to metaphysics. *Philosophy & Technology*, *35*. http://dx.doi.org/10.1007/s13347-022-00508-4

Grush, Rick, & Springle, Alison. (2019). Agency, perception, space and subjectivity.

Phenomenology and the Cognitive Sciences, *18*, 799–818. http://dx.doi.org/10.1007/s11097-018-9582-y

Heidegger, Martin. (1990). *Being and time*.

Kirsh, David, & Maglio, Paul. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.

Menary, Richard. (2012). Cognitive practices and cognitive character. *Philosophical Explorations*, *15*, 147–164. http://dx.doi.org/10.1080/13869795.2012.677851

Merleau-Ponty, Maurice. (2002). *Phenomenology of perception*.

Smart, Paul R., Andrada, Gloria, & Clowes, Robert W. (2022). Phenomenal transparency and the extended mind. *Synthese*, *200*. http://dx.doi.org/10.1007/s11229-022-03824-6

Thompson, Evan, & Stapleton, Mog. (2009). Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi*, *28*, 23–30. http://dx.doi.org/10.1007/s11245-008-9043-2

Wheeler, Michael. (2005). *Reconstructing the cognitive world: The next step*. MIT Press. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=138503

Wheeler, Michael. (2019). The reappearing tool: Transparency, smart technology, and the extended mind. *AI & SOCIETY*, *34*, 857–866. http://dx.doi.org/10.1007/s00146-018-0824-x