

ORIGINAL ARTICLE

Return of the evil genius

Doug Hardman 

Department of Psychology, Bournemouth University, Poole, UK

Correspondence

Doug Hardman, Department of Psychology, Bournemouth University, Poole House, Talbot Campus, Fern Barrow, Poole, BH12 5BB, UK.

Email: dihardman@bournemouth.ac.uk

Abstract

In this essay, I consider whether it makes sense to say that our cognitive capacities—remembering, imagining, intending, hoping, expecting and so on—manifest as inner, subpersonal processes. Given whether something makes sense is a grammatical rather than theoretical or empirical issue, it cannot be explained but can only be better understood by describing and reflecting on situations in which it arises. As such, I approach this issue using the descriptive method of O.K. Bouwsma, which is a development of Wittgenstein's latter methodological approach of conceptually clarifying our bounds of sense. In the course of my investigation, I come to the realisation that cognitive capacities do not, as much psychology and cognitive science imply, make sense as inner, subpersonal processes. Instead, they make sense as personal capacities, which manifest in many ways of acting.

There was once an evil genius, no less powerful than a supremely good god. Tantalising our mother with forbidden fruit. Boasting that he can convince us about anything. Let us assume that this evil genius has returned, only this time it is not just anything about which he hopes to convince us, but our concept of mind. This whispering rabble-rouser means to convince us that our cognitive capacities—remembering, imagining, intending, hoping, expecting and so on—manifest as inner, subpersonal processes.¹ Some claim indeed! I intend

¹Many theories in philosophy of mind or cognitive science—for example, functionalism, eliminativism, naturalistic dualism and even some forms of externalism—share the assumption that cognitive capacities manifest (at least in part) as inner, subpersonal processes, even if they differ markedly on what this means. For discussion, see Block (1980), Chalmers (1997) Churchland (2006), Clark (2008), Fodor (2010), Nagel (1989), Putnam (1975), Ryle (2009) and Searle (1983).

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

© 2023 The Authors. *Philosophical Investigations* published by John Wiley & Sons Ltd.

now to examine how such persuasion is to take place. I expect to discover that the evil genius will only convince us, if he does, by bathing our everyday understanding in a fog. Furthermore, my investigation will uncover not one but two confusions that confound the evil genius: (i) the confusion about what we mean when we talk about cognitive capacities; and (ii) the confusion about what can or cannot be sensibly described as convincing.²

Let us then imagine the evil genius, relaxed, meditating on how he might carry out this persuasion. ‘These humans, with their miserly senses. How can I convince them that their cognitive capacities are sub-personal?’ Then it comes to him. ‘Of course! I must convince them that their capacities have a *locus* and that this is (somehow) inside their head.’ He reflects for a moment. ‘This could be more difficult than I thought. For although cognitive capacities quite obviously have a subject, subjects and loci are clearly not the same thing. I will need all my ingenuity for this one.’ And so he set to work.

Imagine a young woman, Arwa, as astute as she is curious, drifting through the park on her lunchbreak. She closes her eyes, feels the soft sun on her face and daydreams. She thinks back to her recent holiday in the lakes. In particular, how the cold water shocked her as she first dived in off the jetty at the end of the house. Although the park is busy, no one is watching as she lays back on the grass. No one, of course, but the evil genius who is ready to put his plan into action. With the swiftness that his excess powers permit, the evil genius swoops down and hustles Arwa back to his hideout, where his devious process can begin. With Arwa safely sedated, the evil genius gets going with what he thinks is a fool-proof plan. It is so simple. He will just capture Arwa’s capacities, turn them into material things and store them in her head. After all, material things have loci and so if capacities are turned into material things, they will too!

The evil genius meticulously lays out all the equipment he will need for the task. Given cognitive capacities (clearly!) have no locus, he sets aside his traditional tools—scalpel, forceps, probe, retractor—for a specialist set: mind opener, attention grabber, thought probe, memory sifter. It makes sense, the evil genius thinks, to go through all the capacities alphabetically. After all, we don’t want to miss any out. With so many capacities and such little space in Arwa’s head, the evil genius is a bit concerned where everything will go. But, as his friend says, all things excellent are as difficult as they are rare. If this sort of thing was easy, everyone would be doing it.

The first step in his dastardly plan is to capture Arwa’s capacities. He starts with attention but quickly notices a problem. With Arwa sedated, she is not

²The style and philosophical approach of this paper are inspired by Bouwsma’s (1965) *Philosophical Essays*. A succinct way of characterising Bouwsma’s approach is in line with Wittgenstein’s (2009) latter descriptive conception of philosophy: that is, as an activity of conceptual clarification of our bounds of sense. Thus, Bouwsma does not aim to generate philosophical knowledge (in the form of discovery) but *understanding* (in the form of realisation).

attending to anything. He sits down and thinks for a while. There is no other option, he will have to wake her up. Even with the evil genius's powers ensuring Arwa feels no pain, she understandably takes a while to gather herself. 'Look, Arwa,' he says, once his patience runs out, 'it's not complicated, I just want you to focus your attention on something – anything.' It is clear that Arwa is focussed on only one thing: the evil genius himself. That will do. He wastes no time in readying his grabber to capture Arwa's attention. Using his ninth sense—he still can't get over the fact that humans only have five!—he grabs Arwa's attention. Excellent. It takes the evil genius much longer than expected to capture all the capacities, but he eventually gets there. On to step two: matching the capacities with suitable material.

To make things easy for himself, the evil genius decides to match Arwa's capacities to human brain matter, with its attendant electrical activity. On one side of the table he carefully corrals the captured capacities, still in alphabetical order. On the other side, he lays out the material he will match them with. As with step one, it takes the evil genius longer than expected to match the material, but he gets there in the end. With the matching complete, he joins all the capacities together in a shape he thinks will fit neatly in Arwa's head, before carefully cracking it open, ready to stuff in the capacities. However, as he peers down into Arwa's head, he notices something interesting: her brain is awash with activity. Looking closely, the evil genius sees that Arwa's active brain matter is identical in material and construction to what he has matched her capacities with. 'Not what I was expecting at all!' He stops and thinks for a while. Given all of Arwa's capacities and what's already in her head, he cannot see how it will fit. He is also running out of time, so he makes a decision. Given what he has created is identical to what is already in Arwa's head, he decides that it will do no harm to take out what is already there to make room for his new material. With a curtain shielding Arwa from his machinations, he makes the switch.

After collecting stray thoughts (there are always bits and bobs left over with this sort of thing), he wipes Arwa's memory of the process, takes her back to the park on her lunchbreak and waits for her inevitable reaction. Nothing. Arwa just lies there, daydreaming, enjoying the sun as she was. Where's the fun in that? The whole point of this is to get a reaction. The evil genius thinks it through. There is no other option, he will have to break cover again and explain to Arwa what has happened. Just imagine her face when she finds out!

After breaking cover, it takes a while (as it did the first time) for Arwa to gather herself. Once she has, the evil genius begins explaining in earnest. 'Arwa' he says, 'you now accept that I am no less powerful than a god, and I am telling you that I have turned all your cognitive capacities into material things and stored them in your head. Of course, they still seem just like the same capacities you had before, but trust me, they are completely different. They are now material things in your head. For example, try to remember

that leap into the water again, the one off the jetty. That is now completely material! What do you think of that!?' But Arwa, as I already mentioned, is as astute as she is curious. 'I certainly do believe that you could have turned my capacities into material things', she replies. 'I can see now that you are no less powerful than a god and so anything should be well within your capability – even making impossibilities possible! But I think that perhaps a simpler explanation makes more sense. Given it seems to me that nothing whatsoever has changed, maybe what you mean by *material things* is just what I mean by *capacities*. So, when you say that all my cognitive capacities have been turned into material things, it doesn't bother me.' This is not going quite as the evil genius had planned.

'First' he replies, 'let me assure you that all your cognitive capacities have most certainly changed and are now completely material.' To make his point, the evil genius cracks open Arwa's head and shows her the new material he has created. But it looks to Arwa, in all important respects, just like her brain. The evil genius closes her up. 'Look', says Arwa, 'I know you've spent a lot of time on this, but I just think we have a misunderstanding. Material terms just aren't the logically appropriate ones by which to adequately describe things like remembering jumping into the lake off a jetty. Remembering jumping into the lake off a jetty just doesn't make sense, to humans at least, solely in material terms. Of course, we might be able to assign a neurophysiological correlate to my remembering, but this is not going to be a remotely adequate description of what's going on. If we try to describe my remembering in solely material terms, we are simply going to destroy the phenomenon.'³ I'm sorry, I know this isn't what you want to hear.' As you can imagine, the evil genius does not take this sort of thing well. Without so much as a by-your-leave, he storms off in a huff, his attempt at persuasion dissipating in his wake. He needs another, much more devious, plan.

'Right' thinks the evil genius, 'if humans are so ill-equipped to deal with simple things such as capacities being turned into material, I will have to do something much more cunning. I will have to turn their capacities into something *non-material* and locate them inside their heads.' With this new objective in mind, the evil genius sets off to attempt this difficult task. This time, he promises, there will be no mistakes. Given he already knows his way around Arwa's head, he decides that she again will be the first subject to undergo the transformation. As before, he swoops down, steals her from her lunchbreak and prepares the equipment he will need. Again, with Arwa sedated, the evil genius thinks hard about what he will turn her capacities into. Eventually, he alights on a plan. 'Given these humans seem so enamoured with computers, of all things, I will make the relation of their brains and cognitive capacities just

³For more detailed discussion of related issues, see Bennett and Hacker (2022).

like that of computer hardware and software. To that end, I will convert their capacities into symbolic computations, which, like computer software, I can store in the “hardware” of their brain.’ This, the evil genius reflects on modestly, is brilliant.

With his experience from the previous attempt, the evil genius knows that Arwa will need to be awake in order for him to capture her cognitive capacities. Again, he decides to go in alphabetical order, starting with attention. Again, he has the rigmarole of getting Arwa to focus her attention on something. Once he gets going, however, he finds this transformation much easier to effect. Unlike when trying to convert Arwa’s capacities into material things, this time there are no existing symbolic computations clogging up Arwa’s head. In any case, as the computations are nonmaterial, they do not take up any space! The evil genius thus has no problem slotting them in over Arwa’s existing brain matter. After still some not inconsiderable time, the evil genius completes his task. ‘Excellent’, he thinks, ‘this time the persuasion is sure to work.’

But, again, as with his previous attempt, Arwa shows no knowledge of the transformation. There she is, just going about her business as she was: lying flat on the grass, reading her book, smelling the flowers. The evil genius sighs. It looks like he will have to explain it all again. At least this time Arwa will remember him and know from the start that he is no less powerful than a god and thus this is all perfectly well within his capability. However, much like the last time, his attempts at explanation do not go smoothly.

‘Christ’, says Arwa, ‘have you not given up on all this yet! What have you changed my capacities into now? Symbolic computations!?! Sorry, you’ve totally lost me on this one, how is that supposed to work?’

‘It’s perfectly simple’, the evil genius replies, ‘even humans such as you should understand it. To make it easy, I’ll give you an example. Let’s say you remember the lyrics to a song. In this instance your capacity of remembering manifests as an inner computation or representation, in which is stored the relevant information.’

Arwa thinks for a minute. ‘That makes no sense’, she replies. ‘In this instance, my capacity of remembering manifests through the action of actually singing the song. There is no need to posit some vague, sub-personal mental event that precedes this action. In fact, doing so merely pushes back that which we can sensibly account for as the actions of a person, onto subconscious elements for which no such sensible account is available. The talk of an internal mental computation or event adds nothing to the description of remembering which, in this instance, manifests in the action of singing the song. This confusion is even more pronounced’, she continues, pointing across the street, ‘if we consider the example of what that guy’s doing over there: “remembering” how to ride a bike. In this purported instance of “implicit” or “procedural” memory, there is not even conscious awareness of the remembering. He just gets on his bike and rides it. In this instance, therefore, the remembering manifests

merely in the action of him successfully riding his bike – as he is doing – without falling over! Again, attempting to insert some unexplainable preceding mental event is not going to add anything to our extant description of this manifestation of remembering.’

By this point, the evil genius was getting exasperated. ‘Ok Arwa’, he says, ‘given you’ve got all the answers, just try to remember something you did in the past; say, jumping into the lake off that jetty on holiday a few years ago. Try to account for *that* without your capacity of remembering manifesting through creating a mental picture in your head, as an inner symbolic computation!’

‘I certainly don’t deny that such acts of episodic remembering occur through creating a mental picture,’ Arwa replies. ‘But “creating a mental picture” should not be confused with “an inner symbolic computation occurs inside my head”. The creation of the mental picture is the action I, *as a person*, conduct (much like singing a song, or riding a bike). Except colloquially, it makes no sense to say that it occurs “in my head” any more than it makes sense to say that I ride a bike in my head. Look, remembering is, first and foremost, a personal capacity. It is from such an aspect that remembering makes sense. There are many ways of acting through which one’s capacity of remembering can manifest, such as singing a song, riding a bike, or creating a mental picture. But none of these actions make sense as inner, subpersonal events or processes, and our descriptions gain nothing by erroneously inserting such things as preceding the (comprehensible) actions.⁴ You seem to be saying that the actions I have described cannot be all there is to remembering, and that the *real* remembering must be located in some amorphous subpersonal processes. But not only is there no empirical evidence for these processes, they do not even make conceptual sense in the first place.”⁵

By now the evil genius had had enough. He had failed twice in his attempt to convince. He would not try a third time. Before he could again storm off in a huff, however, Arwa had one last point to make. ‘Look, if your intention was to convince, then you have, I’m afraid, failed at the first hurdle: learning and respecting the language of your subjects. For you may well have many excess powers and senses by which cognitive capacities can be turned into, and understood as, material or symbolic computations. But I, alas, can only see, smell, hear, touch and taste. I cannot also *cerpicio!*’⁶

In this essay, I have tried to understand how an evil genius could convince us that our cognitive capacities—remembering, imagining, intending, hoping,

⁴As Hacker (2013) notes, this does not amount to behaviourism, insofar as cognitive capacities are not *reducible to* behaviour.

⁵This account broadly accords with the later Wittgensteinian view on memory, which is explicated well by Moyal-Sharrock (2009, 2013). It also accords with some, although not all, enactivist accounts in cognitive science; see Thompson (2010) and Varela et al. (2016) for discussion.

⁶Cerpicio, of course, being the famous extraterrestrial sense Bouwsma (1965) introduced in his essay ‘Descartes’ Evil Genius’.

expecting and so on—can manifest as inner, subpersonal processes. In so doing, I have tried to foreground what such persuasion would entail and what this means for our understanding of cognitive capacities. In the first case of converting cognitive capacities to material, the evil genius was confronted with the fact that cognitive capacities already make sense in material terms, without such an outrageous transformation. Each manifestation of a cognitive capacity already has a perfectly understandable neurophysiological correlate, even if we cannot identify exactly what this is in each instance. The problem is not that such an aspect makes no sense, but that describing cognitive capacities solely from such an aspect does not in any way adequately account for them. In the second case of converting cognitive capacities into inner symbolic computations, the evil genius was confronted with a different problem. Namely, it is not that there is an existing, perfectly understandable “inner symbolic computations” aspect available. Rather, it is just that, for humans at least, such an aspect is simply not necessary for describing and understanding cognitive capacities. As such, invoking such an aspect to account for the purported transformation makes no sense.

I finish by reflecting not just on how we make sense of cognitive capacities but why the evil genius failed to convince. I hope that in so doing, each problem can shed light on the other. In both cases, the evil genius required special powers and senses in order to “capture” cognitive capacities and “convert” them. If we admit such special powers and senses, then both cases make sense. But, as Arwa implied, humans do not make sense of the phrase “cognitive capacities” (or anything else) by relation to powers and senses only evil geniuses possess. We make sense of them by relation to the powers and senses we have. In both cases, therefore, the evil genius has not only become confused in his understanding of cognitive capacities but has failed in his quest to convince; simply, because the supposed transformations have absolutely no practical effect at all on Arwa (or anyone else). They are merely nonsensical distinctions without a difference.⁷

ORCID

Doug Hardman  <https://orcid.org/0000-0001-6717-2323>

REFERENCES

- Bennett, M. R., and P. M. S. Hacker. 2022. *Philosophical Foundations of Neuroscience*. 2nd ed. USA: Wiley Blackwell.
- Block, Ned. 1980. ‘Troubles with Functionalism’. In *Readings in Philosophy of Psychology*, edited by Ned Block, 268–305. Cambridge, MA: Harvard University Press.
- Bouwsmma, O.K. 1965. *Philosophical Essays*. USA: University of Nebraska Press.
- Chalmers, David J. 1997. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Churchland, Patricia Smith. 2006. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge, Mass.: MIT Press.

⁷I thank the associate editor for the particular turn of phrase ‘a distinction without a difference’.

- Clark, Andy. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.
- Fodor, Jerry A. 2010. *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Hacker, P. M. S. 2013. 'Prologue: Wittgenstein's Philosophy of Psychology as a Critical Instrument for the Psychological Sciences'. In *A Wittgensteinian Perspective on the Use of Conceptual Analysis in Psychology*, edited by Timothy P. Racine and Katherine L. Slaney, 10–27. London: Palgrave Macmillan.
- Moyal-Sharrock, Danièle. 2009. 'Wittgenstein and the Memory Debate'. *New Ideas in Psychology* 27 (2): 213–27.
- Moyal-Sharrock, Danièle. 2013. 'Wittgenstein's Razor: The Cutting Edge of Enactivism'. *American Philosophical Quarterly* 50 (3): 263–79.
- Nagel, Thomas. 1989. *The View from Nowhere*. London: Oxford University Press.
- Putnam, Hilary. 1975. *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Ryle, Gilbert. 2009. *The Concept of Mind*. London: Penguin Classics.
- Searle, John R. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Thompson, Evan. 2010. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. USA: Harvard University Press.
- Varela, F.J., E. Thompson, and E. Rosch. 2016. *The Embodied Mind: Cognitive Science and Human Experience*. London: The MIT Press.
- Wittgenstein, Ludwig. 2009. *Philosophical Investigations*. 4th ed. Chichester: Blackwell.

How to cite this article: Hardman, Doug. 2024. "Return of the Evil Genius." *Philosophical Investigations* 47 (1): 24–31. <https://doi.org/10.1111/phin.12409>.