How phenomenology can contribute to an improved cognitive ontology

Name: Tanja Mourachova
Wordcount: 4000
Date: July 16th, 2024

Table of Contents

Introduction

For many researchers the main aim of cognitive neuroscience is to find the neural mechanisms in the brain that realize cognitive functions. This goal comes with many challenges and one of them is that the cognitive concepts, the explananda that drive the design of experiments and theories, might not be precise enough or might not even refer to existing phenomena at all. Many researchers in the cognitive sciences therefore agree that our cognitive ontology, the taxonomy of the mind, so to speak, is in need of revision. In this paper I will discuss what the desiretata are for this improved cognitive ontology and then I will introduce some results from phenomenology that fit those criteria. My thesis is that phenomenological research can contribute to an improved cognitive ontology.

Section 1: The need for an improved cognitive ontology

1.1 What is cognitive ontology?

The word ‘ontology’ can have different specific meanings in different scientific fields. In this paper I use ‘cognitive ontology’ to refer to a more or less formal set of concepts and categories, a taxonomy, regarding cognition and psychological phenomena, including the relations between these concepts and categories. In daily life we use many concepts to refer to mental phenomena, like ‘belief’, ‘desire’ and ‘will’. This is referred to as folk psychology in philosophical literature and is an informal kind of cognitive ontology. Taking these concepts as a starting point, the cognitive sciences have, through years of experimenting, researching and theorizing about what produces behavior, developed refined concepts that refer to specific psychological phenomena (like ‘motivation’, ‘object perception’, and ‘action planning’) to be able to provide explanations. These concepts are part of a more formal cognitive ontology, a set of concepts that repeatedly are explained and discussed in scientific works, often further broken down into processes involving smaller, more specific, phenomena and concepts that refer to these phenomena. A formal cognitive ontology is very important for doing research, as these concepts are used to set up experiments to measure for example the brain activity that is involved in a psychological phenomenon to which such a concept refers to.

1.2 What is wrong with the current cognitive ontology?

In recent years there has been a growing concern in the neuroscientific field regarding the cognitive ontology that is currently used. The holy grail of neuroscience is to find the neural correlates of all these cognitive concepts used in for example cognitive psychology, such to provide the exact neural mechanisms that underlie the psychological phenomena that are studied. However, attaining this goal has proven to be problematic. In the ideal situation, neuroscientific studies would produce one-on-one mappings between specific concepts and neural structures in the brain. This would show which specific structure or part of the brain is responsible for executing a specific cognitive function (for example “the amygdala is the fear center of the brain”). The problem is that, contrary to what such phrases as above imply, it actually seems to be impossible to find one-on-one mappings. What neuroscientific studies show instead is that many different brain areas or networks light up when a person is showing behavior that is associated with a specific cognitive concept. This is also the case the other way around, specific areas of the brain are involved in a whole array of different cognitive functions.

Pondering on this problem, many scientists have proposed different solutions. Proposed is for example to map neural networks, or possibly other kinds of structures in the brain, instead of areas: the brain can be cut up in different ways, after all (Poldrack 2010, 755). Others have casted doubt on the cognitive concepts themselves, since psychological categories are also contingent, even more so. A beautiful example of this can be found in *Naming the Mind*, a book by Kurt Danziger, a professor of psychology that worked at a university in Asia and came up with the idea to design a course together with a local colleague that also taught psychology. This turned out practically impossible, as they used completely different concepts that referred to psychological phenomena that the other was not used to or aware of (1997, 1-3). In his book, Danziger goes on to describe how the concepts of modern western psychology came to be in the 20th century. The conclusion one can take from his research is that the way the conceptual cake of psychological phenomena is cut, is contingent and depends a lot on the culture and moment in history where the formal or informal ontology is developed. Building on this insight many neuroscientists place the problem of the failed one-on-one mappings on faulty concepts and categories. A proper cognitive ontology with the right categories might solve the problem.

Lastly, some researchers think that the goal to find one-on-one mappings itself is misguided. It is very well possible that the brain uses the same structures for different cognitive functions, and that many-to-one or many-to-many mappings are the only relations that we will be able to

find (Anderson 2015, 74). In any case, the consensus seems to be that the cognitive ontology employed in the cognitive sciences is in need of revision (McCaffrey and Wright 2023, 428).

Section 2: What does an improved cognitive ontology look like?

2.1 Distinct concepts

The first point that often is made is that there should be consistency in the terms and concepts used to describe cognitive processes and which behavioral tasks actually tap into these processes. For example, McCaffrey and Wright refer in their paper (2023) to a study done by Quesque and Rosetti (2020) where they show that social cognition researchers use a motley of terms such as “mindreading,” “mentalizing,” “theory of mind,” “perspective taking,” “cognitive empathy,” and so on to denote the same hypothetical construct. This is also the case the other way around, one concept can refer to many different things. The same study shows that the concept ‘empathy’ has more than 40 different definitions in different papers. To make cognitive science more effective, an improved ontology will have distinct and precise concepts which clearly distinguish phenomena and terms that refer to all these phenomena without overlap.

2.2 Reflecting the brain’s native ontology

This leads us to the next requirement. If the current concepts are going to be changed, splitted-up, fused or discarded, and these revisions are contingent, how to determine which new concepts and categories would be “correct”? In his influential paper, *Mining the Brain for a New Taxonomy of the Mind*, Michael Anderson writes:

> Everyone in the psychological sciences is united in the assumption that the mind is organized, but quite naturally there have been and continue to be debates about the best way for this organization to be described. (...) researchers in the cognitive neurosciences have begun to explore the brain’s native ontology – the categories that it uses to interpret the world – and to use these explorations to motivate revisions to the basic categories of psychology. (Anderson 2015, 68-69)

The “right taxonomy” for understanding cognition should reflect the natural divisions in the organization of the mind, according to Anderson. In other words, the categories should not only exist in the minds of cognitive scientists, but should be really out there in the world, specifically in the brain. These categories or divisions do not have to be physical, as the quote suggests, if

the mind and brain would be compared to a program running on a computer, the software could use categories that are not represented on a hardware level. The next requirement of an improved cognitive ontology is thus that it reflects the native categories used by the brain.

2.3 Unifying the cognitive sciences

Both of the above requirements are motivated by the goal to unify the cognitive sciences (Anderson 2015, 70; McCaffrey and Wright 2022, 449; Piccinini and Craver 2011, 283), which is closely related to the debate about the independence of psychology as a field. Both Anderson as Piccinini and Craver argue that the concepts designed in psychology should be constrained by neuroscientific results. The idea is, like Piccinini and Crave explain, that all the cognitive sciences are united in the aim of explaining behavior and psychological phenomena by finding the mechanisms that underlie cognitive functions. The task of psychology is to find the rough functional distinctions of cognition, in turn neuroscientific research tests these distinctions in experiments to find out where or how the brain implements these functions. The task of neuroscience is to fill the "sketch" that cognitive psychology provides, with more concrete details or feedback on what works or not. Based on this new information cognitive psychology should adjust the structure of the sketch to make the neuroscientific results fit. This iterative process could be compared to how a scientist or philosopher would compose a paper, first drafting a rough structure that will then develop and change throughout the process of writing up concrete points with examples, evidence and side arguments, finally resulting in a cohesive main argument. This way both cognitive psychology and neuroscience can work together to produce a complete multilevel mechanistic explanation of neural systems and the cognitive functions they realize.

2.4 Diversity

Not everybody agrees that a unified cognitive ontology is the way to go. In their paper *Neuroscience and Cognitive Ontology: A Case for Pluralism*, McCaffrey and Wright cite many authors that doubt that it is desirable to have one cognitive ontology. Their arguments is as follows:

> Recent philosophy of science suggests that while developing taxonomies in biology, chemistry, psychiatry, and so on requires correctly tracking the world's metaphysical structure, understanding that structure does not determine what taxonomy we should have. Instead, scientific ontologies unavoidably depend on researchers' goals and interests. (McCaffrey and Wright 2022, 452)

The aim to determine the categories of the mind, as nature has carved them, supposedly in the brain, is misguided, according to McCaffrey and Wright. Taxonomies should reflect the interests and goals of researchers and should not be fixed by understanding the metaphysical structure of the explananda (McCaffrey and Wright 2022, 449). They suggest that the confusion of terms and concepts right now in the cognitive sciences is due to the fact that scientists already do employ different taxonomies depending on their research goals, but are not explicit about it. A pluralistic approach to cognitive ontology, where a multitude of taxonomies is deliberate and coordinated, should take away the hindrance of overlapping concepts experienced now in cognitive neuroscience.

The pluralistic approach to cognitive ontology seems to be incompatible with the second requirement I discussed, namely that the cognitive ontology should reflect the brain's native distinctions. One of the arguments McCaffrey and Wright bring forth is that even if we would be able to visualize brain research data in such a way that we get independent categories, they will be meaningless (by design). The problem is that you can not do anything with meaningless categories, as the link to everyday life is broken and this type of knowledge does not provide anything useful for human life (McCaffrey and Wright 2022, 446).

I am inclined to agree with McCaffrey and Wright, although I think we should, especially in a pluralistic approach, allow the goal to find the brain's native categories to exist besides other research goals. It can still yield interesting results that could inform us about the "world's metaphysical structure", which, according to McCaffrey and Wright themselves, is useful to understand. I will therefore take McCaffrey and Wright's pluralistic approach as a general requirement to the overall cognitive ontology and reduce the requirement to reflect the brain's native ontology to apply only to the specific cognitive ontology or taxonomy that fits that research goal.

To conclude this section, an improved cognitive ontology is thus a set of taxonomies that are geared towards a specific research goal. These taxonomies contain precise and distinct concepts that do not overlap with each other inside the individual taxonomies. Some of the taxonomies are meant to reflect the brain's native ontology, while others do not have to. Researchers should be explicit to what research goal they contribute and which taxonomy they use for this. Lastly, an increased cooperation is desired between different fields within the cognitive sciences. There might be different taxonomies, but this does not mean that each field has their own. The idea is that each taxonomy can be a result of cooperation by multiple fields.

Section 3: Phenomenologically constructed cognitive concepts

3.1 What is phenomenology?

Phenomenology is a branch of philosophy that studies the structure of experience. The idea is to examine phenomena in a direct, unmediated, way from the first-personal perspective and to describe these phenomena the way they appear to us in experience. Many famous philosophers are associated with the phenomenological tradition, notably Husserl, Heidegger, Merleau-Ponty and Sartre. The phenomenological method consists of first bracketing any assumptions about the (existence) of the natural world around us (even scientifically accepted ones) and then turning our attention to our conscious experience. In recent decades there has been more cooperation between phenomenology and the cognitive sciences, as there is a great overlap in phenomena both study.

3.2 Case-study: Ratcliffe's phenomenological distinctions of verbal hallucinations

In his book *Real Hallucinations*, Matthew Ratcliffe discusses anomalous experiences from a phenomenological point of view. Currently the categorization used in scientific literature regarding verbal hallucinations consist of two concepts: thought insertion (TI) and auditory verbal hallucinations (AVH). TI refers to the phenomenon when one experiences their own thoughts as alien. The thoughts appear inside the boundaries of one's subjectivity (there is a felt ownership of the thoughts) but without the sense of agency, as if someone else planted these thoughts in one's head. AVH, on the other hand, refers to the phenomenon of hearing voices in the absence of a speaker. The voices have auditory properties that are perceived as external to the subject, as if there is a real voice speaking but no one else, besides the person experiencing AVHs, can hear it. Both TI and AVH are associated with psychological disorders like for example schizophrenia. Here we see cognitive ontology as a network of connected concepts in action: bigger constructs like 'schizophrenia' are associated with more concrete concepts like 'TI', which in turn is negatively associated with a more specific concept such as 'agency'. After exploring verbal hallucinations with the phenomenological method using first-personal accounts, Ratcliffe concludes that the categories TI and AVH are misguided, which brings into question also the other associated constructs.

One of the reasons why Ratcliffe comes to this conclusion is that, phenomenologically speaking, the distinction between an external voice and an internal thought is not that clear. Sometimes we think quietly, sometimes we say things “out loud” in our minds (we can read in both ways for example). Sometimes we simulate a conversation in our minds where we “hear” this other person speaking in their specific voice, tone and accent. Or when we have a music piece stuck in our head, we “hear” it playing over and over in our minds. There is a kind of auditory quality to these experiences, but not quite the same as real auditory perception. Keeping this in mind when analyzing first-personal accounts of verbal hallucinations, the question arises what people actually refer to when they describe “hearing” a voice. Patients have no other choice but to use words that refer to either auditory perception or imagination, when trying to describe an experience that might be strange and different from what these words technically refer to. It could be that their experience of hearing voices is just as clear as real perception, but it could also be that they are referring to an experience more akin to the experiences I described above. Ratcliffe asserts that, in psychiatric practice, often first-person accounts are misinterpreted, where descriptions are taken too literally. Most AVHs are actually experienced as internal and without real auditory properties, which makes them TIs in disguise (Ratcliffe 2017, 54-62).

After casting doubt on the conceptual distinction between TI and AVH and their definition, Ratcliffe continues to use his phenomenological analysis to question the currently most influential mechanistic explanation of how verbal hallucinations in general arise, namely due to a failure in the source monitoring of generated inner speech. The idea is that the mechanisms involved in thought production resemble those at work in bodily action and motor control, particularly the mechanisms that enable us to experience bodily movements as self-produced versus caused by external forces. Ratcliffe reacts specifically to the account by Christopher Donald Frith (1992), who introduces this connection: “a thought is generated, a signal is produced that predicts its occurrence. A comparator then matches the actual output with what was anticipated” (Ratcliffe 2017, 77). If the occurrence-predicting-signal is not produced, or if there is a mismatch between what was anticipated and what actually happened, there is no experienced “intention” to think that thought and so the thought appears to be non-self-generated and the patient experiences a lack of agency, as if someone else is controlling that thought.

Ratcliffe has many objections to Frith’s explanation, one of them is that a source monitoring mechanism for thoughts appears to be redundant. While we need to distinguish self-generated

actions from other bodily movements, we do not need to distinguish our own thoughts from non-self-generated thoughts, as we never are in a situation where our thoughts are externally generated. A mechanism to distinguish this would thus never evolve (Ratcliffe 2017, 78). Other issues are that this mechanism does not explain the specific content of what is said by “the voice”, its thematic consistency and why it is often attributed to a specific person. Then there are also many other concepts involved in current explanations that rather complicate the issue. All the talk about agency, content, degree of effort, intention, surprise, does nothing to illuminate the phenomenology of VHs, writes Ratcliffe, the “main identifier of voice-hearing (...) is simply the feeling that they do not feel as if they came from me” (Ratcliffe 2017, 70).

Furthermore, a phenomenological examination of first-personal accounts of voice-hearers also shows that instead of lacking anticipation, many patients do anticipate their voices, and anxiously so. Ratcliffe suggests that it is actually this anxious anticipation that causes verbal hallucinations. Keeping in mind that most people that experience “voices” have suffered interpersonal trauma, Ratcliffe proposes the following alternative explanation. When a patient internally simulates (past) conversations with their abuser, in the process of thought crystallization one anticipates an overwhelming negative content, and, unable to escape it, starts to dissociate from it, such that he ultimately experiences it as alien (Ratcliffe 2017, 89). A similar cause could be postulated with verbal hallucinations that have strong auditory properties and that seem to be located externally. It could be very well that a person anxious and hypervigilant about negative affirmations, is more prone to interpret external noises as such (Ratcliffe 2017, 101).

In his book, Ratcliffe invites neuroscientists to use his phenomenologically derived categories (internal non-auditory verbal hallucinations and external auditory verbal hallucinations) to search for the distinct mechanisms that underlie these two different groups of phenomena, both of which seem to involve affective anticipation in a different way (Ratcliffe 2017, 71-72). He also emphasizes the need for cooperation between phenomenological and neurobiological research:

> one will get nowhere regarding VHs as a unitary kind and failing to distinguish hypervigilance VHs from internal VHs, or inner speech VHs from memory VHs. So phenomenology can assist neurobiological research by clarifying *what it is* that requires explanation. (Ratcliffe 2017, 185)

Phenomenology can be of great help to sharpen up the explananda for scientific study, making the borders of phenomena and the underlying mechanisms clearer.

3.3 Would phenomenological concepts improve our cognitive ontology?

In the above case study I have shown that Ratcliffe's phenomenological analysis of verbal hallucinations questions existing cognitive concepts and introduces new distinctions and categories. I could not do justice to the amount of nuance found in his book and all the different cases Ratcliffe considers, however, it is clear that his phenomenological research equipped him with many insights to draw distinctions between for example mute thoughts, thoughts with auditory properties and auditory perception. These more precise distinctions led to the insight that many phenomena, now interpreted as AVH fit the definition of TI better, and that it would make more sense, given the different processes involved, to categorize verbal hallucinations as internal non-auditory and external auditory, and drop explanations that involve other vague concepts such as 'agency', 'ownership' and 'intention'. This kind of revision seems to be exactly what I identified as the first criterion for an improved cognitive ontology: precise and distinct concepts.

What about the second criterion? Do phenomenologically derived concepts reflect the brain's native divisions? In the case of Ratcliffe, it seems to me that he believes that the categories phenomenologists can develop, reflect distinct mechanisms in the brain. The assumption behind this is that careful examination of experience will uncover the natural structure of not only the phenomena themselves, but also the borders of their physical, mechanical or computational constitution. There are however also objections that could be brought up. One could for example object that a conscious bracketing of assumptions does not necessarily make a phenomenologist unbiased. It could still be the case that their prior knowledge, language, or past experiences influences the way they interpret experience and first-personal accounts, and that this influence is inescapable no matter how good the bracketing method is. The question thus remains to which specific taxonomy, to what research goal, phenomenology specifically contributes, but it definitely can have a place in the pluralistic approach.

Lastly, although it is true that not all phenomenologists have the desire to cooperate with the sciences, there is a growing group that does. Ratcliffe builds on scientific knowledge and invites scientists to collaborate and use his insights. This makes the developing taxonomy interdisciplinary, which is the third criterion for an improved cognitive ontology. Thus, to conclude, it seems to me that phenomenological insights can be of great help in improving

cognitive ontology. It meets all the criteria I listed and it is especially valuable by bringing in more distinctions and new categories to the table.

Conclusion

We have seen that a precise formal cognitive ontology is crucial for scientific research, especially for neuroscientific research that aims to map our cognitive concepts to specific mechanisms in the brain. There is discussion on if this goal is attainable, but the consensus is that our cognitive ontology should be improved either way. Taking McCaffrey's and Wright's pluralistic approach, an improved cognitive ontology should contain multiple taxonomies that are geared towards different research goals, and which contain precise and distinct concepts that do not overlap with each other. Another desiderata most researchers agree on, is that these taxonomies should be a product of cooperation between different fields within the cognitive sciences.

I discussed, as a case study, Ratcliffe's phenomenological analysis of verbal hallucinations, to show that phenomenological insights can help develop such distinct and precise concepts. The kind of conceptual revision Ratcliffe proposes seems to be exactly what is needed for an improved cognitive ontology and it meets all the other criteria as well. Therefore, I conclude that cooperation between the cognitive sciences and phenomenological research is a fruitful avenue in the quest to improve our cognitive ontology.

Bibliography

Anderson, Michael L. 2015. “Mining the Brain for a New Taxonomy of the Mind.” *Philosophy Compass* 10 (1): 68-77. https://doi.org/10.1111/phc3.12155.

Danziger, Kurt. 1997. *Naming the mind: How psychology found its language.* London: SAGE Publications.

Frith, Christopher Donald. 1992. *The Cognitive Neuropsychology of Schizophrenia*. London: Psychology Press. https://doi.org/10.4324/9781315785011.

McCaffrey, Joseph and Jessey Wright. 2022. "Neuroscience and Cognitive Ontology: A Case for Pluralism." In *Neuroscience and Philosophy*, edited by Felipe De Brigard and Walter Sinnott-Armstrong, 427-466. Cambridge, Massachusetts: The MIT Press. https://doi.org/10.7551/mitpress/12611.003.0019.

Piccinini, Gualtiero and Carl Craver. 2011. “Integrating Psychology and Neuroscience: Functional Analyses as Mechanisms Sketches.” *Synthese* 183 (3): 283-311. http://www.jstor.org/stable/41477675.

Poldrack, Russell. A. (2010). “Mapping Mental Function to Brain Structure: How Can Cognitive Neuroimaging Succeed?” *Perspectives on Psychological Science* 5 (6): 753-761. https://doi.org/10.1177/1745691610388777.

Quesque, François and Yves Rossetti. 2020. “What do theory-of-mind tasks actually measure? Theory and practice.” *Perspectives on Psychological Science* 15 (2): 384–396. https://doi.org/10.1177/174569161989660.

Ratcliffe, Matthew. 2017. *Real Hallucinations: Psychiatric Illness, Intentionality, and the Interpersonal World.* Cambridge: The MIT Press.