PhiMiSci
Philosophy and the Mind Sciences

Progress and paradigms in the search for the neural correlates of consciousness
Editorial introduction

Sascha Benjamin Fink[a,b,c] (sfink@ovgu.de)
Ying-Tung Lin[d] (linyingtung@nycu.edu.tw)

This article is part of a special issue on "The Neural Correlates of Consciousness", edited by Sascha Benjamin Fink and Ying-Tung Lin.

Twenty years ago, Thomas Metzinger attested that the empirical science of consciousness is in a pre-paradigmatic phase. Some still see the field in this stage. While there are numerous *theories* of consciousness, none of them can claim to be *the* paradigm of the field (even though some — like the Global Workspace Theory, Higher Order Thought Theory, Predictive Processing — are more prominently name checked than others). Pre-paradigmatic stages of science are rare. According to Thomas S. Kuhn (1962), sciences usually progress by going from paradigm to paradigm via scientific revolutions. Before the first paradigm, fields of research are marked by disagreeing writings that try to "build the field anew from its foundations", because scientists are not able to take any "common body of belief for granted" (Kuhn, 1962, p. 13).

How do sciences arrive at a paradigm? One possibility is that they inherit a paradigm from the science they develop out of: Cognitive science established its paradigms after its split from a psychology dominated by behaviourism. But this was not a proper, all encompassing revolution, in two ways. First, some paradigms

[a]Philosophy Departement University of Magdeburg
[b]Center for Behavioral Brain Sciences (CBBS), University of Magdeburg
[c]Humboldt-Universität zu Berlin, Research Training Group 2386 "Extrospection"
[d]Institute of Philosophy of Mind and Cognition, National Yang Ming Chiao Tung University

Fink, S. B. , & Lin, Y.-T. (2021). Progress and paradigms in the search for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *2*, 3. https://doi.org/10.33735/phimisci.2021.

in cognitive science can still be seen as left-overs from behaviouristic psychology. Second, behaviouristic psychology was not overcome: Some psychologists still remain true to behaviourism. The danger of such incomplete revolutions lies in their possible instability: As long as the not-fully-overcome paradigm has its prominent adherents, it may simply take over again. This may not always be a bad thing. But in order to stabilise an incomplete revolution, secession is needed: Cognitive scientists needed to establish a field of their own and defended it against behaviouristic psychology by searching for allies in different fields (linguistics, robotics, cybernetics, computer sciences, neuroscience) and by doing "boundary works" (Gieryn, 1983).

Some see the contemporary science of consciousness as a development out of cognitive science and thereby simply transpose paradigms of cognitive science onto this new field of research. Others resist this conservative effort and propose alternative paradigms. In such a climate of balkanisation, researchers need to find common ground among the disparate fractions in order to lay the foundations for a unified field of research. One way to find common ground is in the abstract: Metatheoretic reflection and reasoning independently of a paradigm (sometimes known as philosophy), may lead to shared starting points. For example, most researchers accept that consciousness exists and needs to be accounted for. Another way to find common ground is in the concrete: the data all theorists need to account for.

Data can provide common ground if it can be framed in a way acceptable to several fractions. In the science of consciousness, one kind of datum to account for might be how we describe our subjective experience of ourselves and the world, i.e., our phenomenology. But how do we arrive at a phenomenology? Due to the subjectivity and privacy of our experiences, the call for first-person methods can be heard in large parts of the field. However, there is reasonable scepticism vis-à-vis first-person methods, stemming in part from the idea that science must uphold objectivity and allow for intersubjective checks. First-person methods might be hardly compatible with that.

But at least something is commonly accepted: that neural processes correlate with our experiences. The search for *neural correlates of consciousness* (NCCs) therefore plays a fundamental role for a pre-paradigmatic science like empirical science of consciousness: It is one of the few sets of largely uncontested data that would-be theory- and paradigm-builders have to take into account. Why does *this* neural activation correlate with *that* experience?

The second fundamental reason why neural correlates have proven to be so central to this field is their metaphysical promiscuity: They are not committed to one specific theory on how consciousness and the brain actually relate. Still, whichever metaphysics philosophers may converge on — identity theory? epiphenomenalism? supervenience? grounding? neutral monism? pre-established harmony? interactionist dualism? —, as long as consciousness *and* the brain exist, we can expect some statistical coupling between neural and phenomenal variables, i.e., a correlation. James Ward (1911, p. 600), who already used the term "neural

Fink, S. B. , & Lin, Y.-T. (2021). Progress and paradigms in the search for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *2*, 3. https://doi.org/10.33735/phimisci.2021.

correlates of consciousness" in the entry on *Psychology* in the 11th edition of the *Encyclopedia Britannica*, called this *methodological dualism,* where we accept at the beginning that we grasp the two "correlated" entities with different methods; but he already speculated that this may lead to a metaphysical monism in the long run. This commitment to a purely statistical and largely measurable relation between the neural and the phenomenal — correlation — was Crick and Koch's (1990) elegant way of bracketing ongoing metaphysical turf wars thirty years ago: Let them fight, we'll start our science.

The discussion around the NCC therefore combines the abstract and the concrete to foster common ground: The meta-theoretical reflection on a metaphysically neutral but empirically accessible relation between neural and phenomenal variables gave rise to a suitable way to capture data and bracket discussion that could have divided the field prematurely.

The search for NCCs therefore was central at the dawn of the science of consciousness. The question is whether it still is. Maybe the science of consciousness has progressed far enough to overcome NCCs. Maybe the problems associated with NCCs are an indicator of a fundamental crisis this field will face if it binds itself to this concept. Then, we should embrace alternatives, like the neural difference makers of consciousness which Klein, Hohwy, and Bayne (2020) introduced in the first part of this special issue. But maybe we should shy away from calls to leave the search for NCCs behind us. If we abandon the search for NCCs in favour of more elaborate or theory-bound approaches, we may risk fragmenting the young field yet again — simply because it may lead into the usual cycle of paradigms, revolutions, and secessions. But maybe, a convergence onto one and only one paradigm is not even needed in the sciences.

An alternative view, one we find quite plausible, is that the science of consciousness does not mirror the natural sciences in its use of paradigms; rather, it is more like fields such as economics. In the natural sciences, one paradigm governs all — and is largely replaced after a revolution. The usual business of scientists is nearly always done completely under one theoretical roof. In fields like economics, however, paradigms like the Chicago School are not used to establish a normal scientific practice of solving problems arising from that paradigm. Instead, they are reference points used to *orient* scientists in the field. As Paul Hoyningen-Huene and Nicholas Wüthrich (Hoyningen-Huene, 2020; Hoyningen-Huene & Wüthrich, 2017) point out, knowing orientational paradigms allows communicating how one's approach differs from them. One needs to know them to work in this field, but one needs not accept them. In the science of consciousness, theories might fulfil the role of orientational paradigms. However, no theory can yet claim to be such an orientational paradigm, even though some may have the potential. But when there were no empirically adequate theories during the first years of the neuroscience of consciousness, the concept of an NCC fulfilled the role of an orientational paradigm, one that continues to overlap and build

Fink, S. B. , & Lin, Y.-T. (2021). Progress and paradigms in the search for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *2*, 3. https://doi.org/10.33735/phimisci.2021.

common ground for a broad range of theories in the field. The NCC thereby enabled the progress we have seen in this science since the 1990s.

Here, we continue to celebrate the progress made in the science of consciousness, specifically on the NCC, and for a good reason: It has been 30 years since Francis Crick and Christof Koch put the search for NCCs in their *Towards a Neurobiological Theory of Consciousness* (1990) back on the table and 20 years since the conceptual contribution by David Chalmers on the question *What is a Neural Correlate of Consciousness?,* which was published in Thomas Metzinger's seminal anthology on the NCC (2000; see also the introduction by Fink, 2020, to the first instalment of this issue). A lot has happened since these works shaped the field we work in today, reflected in the papers collected here. Most of the authors in this issue specifically address the tension mentioned above, namely between metaphysics, specific theories of consciousness, and the goal of facilitating scientific progress.

Alex Lepauvre and Lucia Melloni (2021) address the methodological challenge NCC research is facing and present three strategies that will accelerate progress in the field. The currently prevalent contrastive method, in which neural activities correlating with conscious states are contrasted with those correlating with non-conscious states, have encountered difficulties in dissociating the proper NCC from neural precursors or successors. Lepauvre and Melloni suggest moving beyond the contrastive method. First, they call for adopting methodologies that capture the phenomenological dimension by incorporating a systematic description of phenomenal consciousness. Second, large-scale coordinations and aggregation of data across paradigms, stimuli, and research groups in common databases must be facilitated. Third, the field needs to adopt pre-registered adversarial collaborations, which directly test contradicting predictions. All of these, they argue, will move the field forward.

Chen Song (2021) underlines the importance of neural structures for the science of consciousness. Her presentation of the structures of the central nervous system illustrates the various ways in which neural structures shape the way in which we experience. While differentiation and integration are considered to be essential to conscious experience, the brain must be able to facilitate these properties. Song argues that the key features of the brain are the structural diversity between neurons and the modular topology of neuronal connections which allow the brain to generate differentiated yet integrated activity patterns. This view is empirically supported by studies that look into inter- and intra-individual changes: Structural diversity is positively correlated with the diversity in conscious experience.

Sascha Benjamin Fink, Lukas Kob, and Holger Lyre (2021) also argue for the view that structures are important in the search for neural correlates of consciousness — so far so that they constrain what can count as a proper neural correlate of an experience (in contrast to a mere statistical correlate). They advocate a *neurophenomenal structuralism* and argue that (i) what type-individuates any ex-

Fink, S. B. , & Lin, Y.-T. (2021). Progress and paradigms in the search for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *2*, 3. https://doi.org/10.33735/phimisci.2021.

perience in a domain (i.e., what makes it specifically of that type) depends on its relative similarity and difference to all other experiences in that domain and (ii) that these individuating structures in the phenomenal domain must be mirrored in their neural substrate in a surjective homomorphism. This *structural similarity constraint* allows progress in so far as it allows for distinguishing merely statistical correlates from those neural events in which consciousness has its physical foothold.

Wanja Wiese and Karl J. Friston (2021) argue for an expansion of the search for correlates: Not only neural correlates should be on the agenda, but also *computational* correlates of consciousness (CCC). Finding such CCCs would allow for a much easier transfer of our empirical findings established in human experimentation to more exotic instances, from non-human animals to artificial systems. Wiese and Friston locate this endeavour within the larger framework of the Free Energy Principle and leverage its mathematics to address not only whether there can be "islands of awareness" (and if so, whether they are marked by complexity reduction or not), but also how to distinguish a system that truly *is* conscious from one that only *simulates* consciousness. Progress in the field can then also come from accepting a minimal model of consciousness or from transgressing from a neural to a computational perspective.

Tobias Schlicht and Krysztof Dolega (2021) doubt that predictive processing (PP), a paradigm currently in high regard, has the power to become a theory of consciousness in its own right or can work as an overarching framework in the neuroscience of consciousness, as suggested by Jakob Hohwy and Anil Seth (2020) in the first part of this special issue. Hohwy and Seth argued that PP might be a unifying framework for the science of consciousness, out of which it may develop into a theory of consciousness in its own right. Schlicht and Dolega strongly disagree: Not only is the evidence in support of PP still equivocal with regard to empirically measurable neural activations, mechanisms, and architectures, but even if this were not the case, the framework's conceptual attachment to consciousness is also too loose. Their article cautions against adopting PP and its allies (e.g., Friston's Free Energy Account, as presented in the contribution of Wiese & Friston in this issue) as overarching models, frameworks, or theories for the neuroscience of consciousness.

Morten Overgaard and Asger Kirkeby-Hinrup (2021) caution against overblowing results from NCC research: Finding the NCC will not answer all of our questions because it is unclear how our theories of consciousness map onto data gathered by NCC research. We are missing an isomorphism between our conceptual frameworks for theories of consciousness and the brain. So even if we knew the NCC, it would be unclear what this says about our theories of consciousness. This situation is worsened by the fact that one of the correlates, phenomenal consciousness, is not directly accessible. The development of theories should take the missing isomorphism between conceptual frameworks as well as theories about consciousness on the one hand and neuroscientific data about

Fink, S. B. , & Lin, Y.-T. (2021). Progress and paradigms in the search for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *2*, 3. https://doi.org/10.33735/phimisci.2021.

NCCs on the other hand into account. They thereby emphasise the distance between NCC research and metaphysical theorising that already motivated Crick and Koch 30 years ago, but which numerous researchers have since attempted to bridge.

Michael Pauen (2021) advocates for a bolder approach: The search for NCCs is not as metaphysically neutral as commonly expected. He considers what it means for the search of NCCs to be "theoretically neutral." Drawing from the history and philosophy of science, in particular the development of the identity between water and H2O, he addresses the importance of the dependency between theoretical assumptions and empirical facts. The author argues that theory dependence is crucial for NCC research and is already present in pain studies.

Editorial note: *Due to Covid-19, numerous steps in the publication of both parts of this special issue were delayed. This foreword has therefore been updated over the course of 2021 to reflect the new additions to this special issue. For their support, we are grateful to all contributors, to all submitters, and to all reviewers. We also want to thank the editorial board, and specifically Wanja Wiese, as well as Fabian Fuchs and Anouk Kinzel, for their support during the typesetting process.*

Fink, S. B. , & Lin, Y.-T. (2021). Progress and paradigms in the search for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *2*, 3. https://doi.org/10.33735/phimisci.2021.

References

Chalmers, D. J. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.), *Neural correlates of consciousness: Empirical and conceptual questions* (pp. 17–39). MIT Press.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275.

Fink, S. B. (2020). A double anniversary for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *1*(II). https://doi.org/10.33735/phimisci.2020.II.85

Fink, S. B., Kob, L., & Lyre, H. (2021). A structural constraint on neural correlates of consciousness. *Philosophy and the Mind Sciences*, *2*, 7. https://doi.org/10.33735/phimisci.2021.79

Gieryn, T. F. (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review*, *48*(6), 781–795. http://www.jstor.org/stable/2095325

Hoyningen-Huene, P. (2020). Orientational paradigms. *Unpublished manuscript.*

Hoyningen-Huene, P., & Wüthrich, N. (2017). Why is general philosophy of science as yet irrelevant for economics and what can be done against this? In *EIPE 20th anniversary conference book of abstract.* https://www.eur.nl/sites/corporate/files/Book_of_Abstracts_.pdf

Klein, C., Hohwy, J., & Bayne, T. (2020). Explanation in the science of consciousness: From the neural correlates of consciousness (NCCs) to the difference makers of consciousness (DMCs). *Philosophy and the Mind Sciences*, *1*(II). https://doi.org/10.33735/phimisci.2020.II.60

Kuhn, T. S. (1962). *The structure of scientific revolutions.* University of Chicago Press.

Lepauvre, A., & Melloni, L. (2021). The search for the neural correlate of consciousness: Progress and challenges. *Philosophy and the Mind Sciences*, *2*, 4. https://doi.org/10.33735/phimisci.2021.87

Overgaard, M. S., & Kirkeby-Hinrup, A. (2021). Finding the neural correlates of consciousness will not solve all our problems. *Philosophy and the Mind Sciences*, *2*, 5. https://doi.org/10.33735/phimisci.2021.37

Pauen, M. (2021). Why NCC research is not theory-neutral*Philosophy and the Mind Sciences*, *2*, 10. https://doi.org/10.33735/phimisci.2021.9188

Schlicht, T., & Dolega, K. (2021). You can't always get what you want: Predictive processing and consciousness. *Philosophy and the Mind Sciences*, *2*, 8. https://doi.org/10.33735/phimisci.2021.80

Song, C. (2021). Brain structural complexity and consciousness. *Philosophy and the Mind Sciences*, *2*, 6. https://doi.org/10.33735/phimisci.2021.9185

Ward, J. (1911). Psychology. In *Encyclopedia Britannica: Vol. XXII* (pp. 547–604). Cambridge University Press.

Wiese, W., & Friston, K. J. (2021). The neural correlates of consciousness under the free energy principle: From computational correlates to computational explanation. *Philosophy and the Mind Sciences*, *2*, 9. https://doi.org/10.33735/phimisci.2021.81

Fink, S. B. , & Lin, Y.-T. (2021). Progress and paradigms in the search for the neural correlates of consciousness: Editorial introduction. *Philosophy and the Mind Sciences*, *2*, 3. https://doi.org/10.33735/phimisci.2021.