CORTEX 175 (2024) A1–A4

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Editorial

Strengthening derivation chains in cognitive neuroscience: Closing editorial

Daniel Mirman [a,*], *Anne Scheel* [b], *Anna-Lena Schubert* [c] *and Robert D. McIntosh* [a]

[a] *University of Edinburgh, Edinburgh, UK*
[b] *Utrecht University, Utrecht, The Netherlands*
[c] *Johannes Gutenberg University of Mainz, Mainz, Germany*

ARTICLE INFO

Article history:
Received 19 April 2024
Accepted 22 April 2024
Action editor Sergio Della Sala
Published online 26 April 2024

1. Introduction

This volume of Cortex sees the publication of the last three of the nine articles forming the special issue, *Strengthening Derivation Chains in Cognitive Neuroscience* (Mirman et al., 2022). Our goal in proposing this topic was to encourage and collect exemplars of rigorous work that lays and strengthens the foundations of research in cognitive neuroscience, from the conceptual entities studied, to the tools and measures to operationalise them and the pipelines through which we process the data. These are core elements of the "derivation chains" that connect abstract scientific claims to concrete observations and thus underpin higher-level hypothesis tests (Meehl, 1990; Scheel et al., 2021). The validity of any hypothesis test (and indeed any inference to abstract claims from data) depends upon the integrity of its derivation chain, and so we need to ensure that the constituent links are secure.

The articles in this collection exemplify this kind of foundational research and illustrate its diversity. They also suggest to us some roads for improving cognitive neuroscience research, which we point towards in this brief editorial. The nine articles will be considered in relation to two main themes: methodological development and concept development.

2. Methodological development

The centrality of measurement and methodological validity is clear: we need to be sure that our research tools are up to job. Measurement and methodological validation studies can be valuable at many different stages of a research program.

If we are developing an entirely new method or measure, then careful validation is essential and cannot be profitably skipped. In the field of meta-science, Isager et al. (2021) proposed a formal method of ranking studies for replication value. This method could help to optimise the targeting of replication attempts and make the selection process more objective. In their special issue contribution, Isager et al. (2024) evaluate the feasibility of this method in practice, identifying several challenges that would need to be overcome before it is ready for wider adoption. These challenges include the clear delineation of topic boundaries, and the quantification of uncertainty in fMRI studies. Future work can build on these insights and further refine, or redefine, this nascent approach to selecting replication targets.

For more established research paths, we may be testing how well existing methods operationalise a construct of interest. The message will not always be reassuring, as Madhavan et al. (2024) found when assessing several common measures of infants' interest in object categories: parental report; overt choices; preferential looking; and pupil dilation.

* *Corresponding author.*
E-mail addresses: dan@danmirman.org, daniel.mirman@ed.ac.uk (D. Mirman).
https://doi.org/10.1016/j.cortex.2024.04.004

The lack of reliable relationships between these measures raises concerns about their validity as measures of a single construct and implies that this core measurement problem needs to be addressed before such methods can be taken further.

In other cases, the focus may be on refining our understanding of what is being reflected by a measure. This may be particularly effective when alternative possibilities can be instantiated in computational models, against which the outcomes can be evaluated, as illustrated by Michaelov and Bergen's (2023) assessment of N400 sensitivity to different aspects of word predictability. Investigating the reliability of our measures may also generate testable insights about the higher-level construct itself. In this vein, Thunberg et al. (2024) have assessed the reliability of multiple behavioural and EEG measures of response inhibition for stop signal tasks. They report a striking pattern of high split-half reliability (within-sessions) and low test-retest reliability (across-sessions), which suggests that response inhibition may have mainly state-like and situational determinants, rather than reflecting a temporally stable, trait-like property of individuals.

Even when we are confident that existing measurement methods are generally valid, there may be work to be done to generalise those methods to new populations (e.g., Turoman et al., 2024), or to optimise the data processing pipelines for different applications. Researchers using an established method are typically aware of the decision points and variations available, and can profitably set up a study (or secondary analysis) to examine the influence of those variations. Examples of this approach are provided by Rodrigues et al. (2024), who explore the influence of methodological and analysis choices for feedback-related negativity in ERPs, and Loenneker et al. (2024), who focus on the preprocessing of reaction time data. Mapping out how different methodological choices influence outcomes is useful in its own right, but Loenneker et al. take this further by developing a checklist for the preprocessing of reaction time data, supported by an expert consensus survey. They thereby crystallise their findings into concrete guidelines that may help to improve and standardise reporting for the field.

As these examples make clear, evaluations of methodological validity can be of value in many ways. It may be useful to frame methodological development itself as a process of incremental accumulation, analogous to research on empirical phenomena. We should not expect a single article to be definitive about a method, just as we do not expect a single article to be definitive about a phenomenon or hypothesis. We should, however, expect that such research is undertaken and published with sufficient visibility to inform the respective scientific community. If method evaluation is not valued or needs to clear a higher bar for publication, we run the risk of accumulating research results that rely on fatally flawed assumptions or idiosyncratic, unstandardised methodological procedures that hinder evidence synthesis. Recognising methodological research as a continuous and incremental process, and providing journal space for it, is crucial for solidifying the foundations of the derivation chains that support our hypotheses tests and enhancing our overall confidence in research findings.

3. Concept development

In cognitive neuroscience, as in science more broadly, we typically study some operationalisation of a construct of interest, rather than the construct directly. Our constructs should be sufficiently well-defined to carve out a coherent set of phenomena, differentiated from other phenomena, and broad enough to be useful but narrow enough for concrete operationalisation and systematic study. An essential step is to evaluate the coherence and appropriateness of our target constructs. This is conceptually challenging, which may help explain why only two papers addressing themselves to this higher level were accepted for this Special Issue. These two articles approach the issue from interestingly different angles.

Katyal and Fleming (2024) articulate a concern regarding the trajectory of metacognition research. They argue that, while developing precise measures for specific facets of metacognition is valuable, it poses a risk of inadvertently narrowing the field's focus. Well-defined measures of the moment-to-moment metacognitive experience of one's own performance can be derived from confidence ratings given during cognitive tasks (Fleming & Lau, 2014). The widespread adoption of these methods, however, risks morphing the field into a relatively narrow science of those measures, and the sorts of tasks to which they are applicable, losing sight of the broader sweep of phenomena that the concept of metacognition originally encompassed (Flavell, 1979). Beyond alerting researchers to this encroaching conceptual myopia, Katyal and Fleming suggest how metacognition research could be expanded to encompass the formation of beliefs about the abilities of oneself and others over longer timescales, and to include neglected aspects such as insight into affective states and traits.

By contrast, Castro et al. (2023) examine aphasia, a neuropsychological disorder of language, known for over 150 years, for which multiple well-established diagnostic tests now exist. The authors conducted a content analysis of six prominent tests of aphasia, to check how well they converged onto a common construct. To their surprise, they found that the different tests had only modest overlap, and substantial differences, suggesting different operational definitions. This sounds an alarm for aphasia research, and calls for efforts to better standardise, and update if necessary, the operational definition of aphasia. Such a project would combine rigorous psychometrics with qualitative development of the target constructs, and involve all interested parties, including scientists, health professionals, people with aphasia, and others. In this case, the discovery of unplanned and unhelpful variations in our measures and constructs may provide an impetus for renewed consensus building, to strengthen conceptual foundations for the future.

4. Making foundational research mainstream

The goal of this special issue was "to stimulate and showcase foundational research that directly targets derivation-chain elements", which is currently underrepresented in the

cognitive neuroscience literature. Making such research more mainstream requires dedicated journal space and editors and reviewers who know how to evaluate it. In editing this special issue, we found that the submissions on methodological development fit well into the mould of regular research articles, being mostly published as *Exploratory Reports* or as standard *Research Reports*. Going forward, the *Registered Reports* article format could also be well-suited to foundational empirical work, with the added benefit of reducing the risk of publication bias. However, the submissions on concept development posed a bigger challenge. Such articles tackle more complex, abstract problems, contain more philosophical elements, and may have a less goal-oriented structure than regular empirical research or reviews, and their outcomes may be more difficult to evaluate. Fostering conceptual derivation-chain research may require more changes to the editorial approach.

As a partial solution to the problem, we suggest taking a more community-based perspective for such higher-level conceptual developments. While it is possible for individual researchers or groups to shape a field's understanding of a concept or method, we would particularly welcome the future creation of consensus-based articles, written by a broad range of experts, and other interested parties. In cases where consensus cannot be achieved, due to theoretical differences (the parties cannot agree on terms) or practical barriers (they are unable to co-ordinate the consensus process), the *Cortex Discussion Forum* offers an alternative approach to work towards it. This format includes a target article and several commentaries, usually capped with a concluding response or synthesis. It provides a platform for the target article to articulate an analysis or proposal from one perspective, and for commentators to pick out points of agreement, elaboration or contention, perhaps leading to resolution and consensus, but at least laying out the issues that need to be addressed.

5. Looking ahead to firmer foundations

Scientists often talk (and teach) as if hypothesis-testing were the main order of business. This approach is portrayed as yielding data that either corroborates or refutes a given theory. But in practice, the link between data and theory is usually too weak to support such inferences. The survey of *Cortex* publications we reported in the opening editorial (Mirman et al., 2022) suggested that when hypothesis-testing studies are published in cognitive neuroscience, their predictions are often rather loose and underspecified, being often based on intuition or previous findings rather than on concrete theories. In such cases, hypothesis tests fall short of their potential because it is unclear which observations would be expected if the tested claims were true or false. More informative tests require strong derivation chains between data and tested claims: a set of conceptual definitions, operationalisations, and auxiliary assumptions that are validated and accepted by the scientific community. These derivation chain elements require dedicated research of their own. We are pleased that this special issue showcases such foundational research, directed at consolidating the concepts we study, sharpening the tools we use, and strengthening the links between the higher-level constructs and their concrete operationalisations. We hope that, as well as being valuable research contributions, these articles serve as exemplars of excellence to stimulate further efforts along similar lines. By hosting this special issue, *Cortex* signals its place as a premier publication venue for such foundational efforts.

CRediT authorship contribution statement

Daniel Mirman: Conceptualization, Writing – original draft, Writing – review & editing. **Anne Scheel:** Conceptualization, Writing – original draft, Writing – review & editing. **Anna-Lena Schubert:** Conceptualization, Writing – original draft, Writing – review & editing. **Robert D. McIntosh:** Conceptualization, Writing – original draft, Writing – review & editing.

REFERENCES

Castro, N., Hula, W. D., & Ashaie, S. A. (2023). Defining aphasia: Content analysis of six aphasia diagnostic batteries. *Cortex, 166*, 19–32. https://doi.org/10.1016/j.cortex.2023.05.005

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience, 8*, 443. https://doi.org/10.3389/fnhum.2014.00443

Isager, P. M., Lakens, D., van Leeuwen, T., & van't Veer, A. E. (2024). Exploring a formal approach to selecting studies for replication: A feasibility study in social neuroscience. *Cortex, 171*, 330–346. https://doi.org/10.1016/j.cortex.2023.10.012

Isager, P. M., van 't Veer, A. E., & Lakens, D. (2021, August 24). Replication value as a function of citation impact and sample size. https://doi.org/10.31222/osf.io/knjea.

Katyal, S., & Fleming, S. M. (2024). The future of metacognition research: Balancing construct breadth with measurement rigor. *Cortex, 171*, 223–234. https://doi.org/10.1016/j.cortex.2023.11.002

Loenneker, H. D., Buchanan, E. M., Martinovici, A., Primbs, M. A., Elsherif, M. M., Baker, B. J., & Pronizius, E. (2024). We don't know what you did last summer. On the importance of transparent reporting of reaction time data pre-processing. *Cortex, 172*, 14–37. https://doi.org/10.1016/j.cortex.2023.11.012

Madhavan, R., Malem, B., Ackermann, L., Mundry, R., & Mani, N. (2024). An examination of measures of young children's interest in natural object categories. *Cortex*. https://doi.org/10.1016/j.cortex.2024.02.015

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66*(1), 195–244. https://doi.org/10.2466/pr0.1990.66.1.195

Michaelov, J. A., & Bergen, B. K. (2023). Ignoring the alternatives: The N400 is sensitive to stimulus preactivation alone. *Cortex, 168*, 82–101. https://doi.org/10.1016/j.cortex.2023.08.001

Mirman, D., Scheel, A. M., Schubert, A.-L., & McIntosh, R. D. (2022). Strengthening derivation chains in cognitive neuroscience: A special issue of Cortex. *Cortex, 146*, A1–A4. https://doi.org/10.1016/j.cortex.2021.12.002

Rodrigues, J., Müller, S., Paelecke, M., Wang, Y., & Hewig, J. (2024). Exploration of the influence of the quantification method and reference scheme on feedback-related negativity and standardized measurement error of feedback-related

negativity amplitudes in a trust game. *Cortex*. https://doi.org/10.1016/j.cortex.2024.02.009

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why hypothesis testers should spend less time testing hypotheses. *Perspectives on Psychological Science, 16*(4), 744–755. https://doi.org/10.1177/1745691620966795

Thunberg, C., Wiker, T., Bundt, C., & Huster, R. J. (2024). On the (un) reliability of common behavioral and electrophysiological measures from the stop signal task: Measures of inhibition lack stability over time. *Cortex*. https://doi.org/10.1016/j.cortex.2024.02.008

Turoman, N., Fiave, P. A., Zahnd, C., deBettencourt, M. T., & Vergauwe, E. (2024). Decoding the content of working memory in school-aged children. *Cortex, 171*, 136–152. https://doi.org/10.1016/j.cortex.2023.10.019