Neuroscience Informatics 5 (2025) 100199

Contents lists available at ScienceDirect

Neuroscience Informatics

journal homepage: www.elsevier.com/locate/neuri

Letter to the Editor

Leveraging transparent ontology learning to refine constructs in neuroscience

David Moreau *, Kristina Wiebels

School of Psychology and Centre for Brain Research, University of Auckland, Auckland, New Zealand

ARTICLE INFO

Article history:
Received 24 March 2025
Accepted 25 March 2025
Available online xxxx

1. Introduction

Psychological constructs, such as attention, impulsivity, and well-being, serve as foundational building blocks for linking brain function to cognition and behavior. Although these constructs are central, their definitions and measurements often lack clarity and consistency. For instance, impulsivity is a construct with diverse and sometimes conflicting definitions across contexts and disciplines: it may refer to a failure to delay gratification, a tendency to act without forethought, or deficits in inhibitory control, depending on whether it is measured through behavioral tasks, self-report questionnaires, or neuropsychological assessments [1–3]. These divergent interpretations lead to conceptual ambiguities that hinder cumulative scientific progress and complicate efforts to map impulsivity to specific neural mechanisms. Here, we explore the potential of ontology learning to advance neuroscience research by addressing challenges in construct measurement and integrating insights across diverse datasets and paradigms. Specifically, we propose the integration of Explainable Artificial Intelligence (XAI) with ontology learning to address gaps in the existing literature. While previous research has explored possible uses of ontology learning in neuroscience, its combination with XAI provides a novel, powerful framework for enhancing transparency, interpretability, and scalability in ontology construction.

2. Ontology learning: concepts and methods

Ontology learning—the process of extracting, organizing, and formalizing knowledge into structured representations, known as ontologies—provides a promising avenue for resolving ambiguities in a wide number of fields, including neuroscience. Initially applied in bioinformatics to unify disparate datasets (e.g., Gene Ontology; Ashburner et al. [4]) and more recently in clinical medicine to standardize terminology (e.g., SNOMED CT; Cornet & de Keizer [5]; Lee et al. [6]), ontology learning is increasingly recognized for its potential to address conceptual challenges in neuroscience [7]. At its core, ontology learning is based on Gruber's [8] definition of ontology as an explicit specification of a conceptualization, involving not only cataloging concepts but also defining their hierarchical and relational structures. The result is a dynamic, machine-readable framework that facilitates data integration, hypothesis generation, and automated reasoning [9].

Ontology construction is a dynamic process that integrates both manual and automated methodologies, each with distinct advantages and limitations. Manual approaches, spearheaded by domain experts, excel in ensuring conceptual accuracy and preserving the nuances of specialized knowledge. However, this process is inherently labor-intensive, time-consuming, and susceptible to individual biases, which can limit scalability and objectivity [4]. Automated methods and tools address these limitations by employing machine learning algorithms capable of identifying patterns and relationships within large datasets, significantly enhancing efficiency. For example, platforms like Protégé (protege.stanford.edu) provide an accessible interface for creating, editing, and visualizing ontologies, streamlining the manual construction process. Formal languages such as the Web Ontology Language (OWL; we3.org/OWL) ensure interoperability across diverse systems and applications. Other tools, such as semantic similarity metrics, quantify the relationships between terms, enhancing the coherence and usability of ontological frameworks [7]. These tools, when integrated effectively, can empower researchers to build robust ontologies that support advancements in neuroscience, as well as in fields such as bioinformatics, artificial intelligence, and medicine.

Yet despite their advantages, automated approaches may struggle to capture the domain-specific subtleties essential for high-quality ontology construction [10]. As a result, hybrid method-

* Corresponding author. Science Building, 23 Symonds Street, Auckland 1010, New Zealand.
E-mail address: d.moreau@auckland.ac.nz (D. Moreau).

https://doi.org/10.1016/j.neuri.2025.100199

ologies, which combine expert input with automated tools, have emerged as a preferred strategy, offering a balance between efficiency and precision. In neuroscience, the lack of standardized definitions for constructs such as "cognitive control" and "working memory" often leads to divergent interpretations and challenges in synthesizing research findings across studies and domains [11,12]. These inconsistencies hinder cumulative progress and limit the generalizability of research findings, as construct definitions and their associated neural correlates remain ambiguous or context-dependent [13]. Ontology learning provides a systematic approach to addressing these challenges by formalizing the relationships among constructs, tasks, and outcomes, offering a more precise conceptual vocabulary for neuroscience. This process enables the identification and mapping of latent structures within datasets, promoting a shared framework that facilitates replication and cross-study comparisons [14]. Such efforts are particularly crucial given the growing reliance on large-scale, data-driven studies, where the risk of conflating distinct constructs is amplified without clear ontological frameworks [15].

3. Using XAI for transparent, interpretable ontology learning

Via its ability to enhance the interpretability and transparency of both the data inputs and the resulting ontological structures, XAI offers a complementary approach to address these challenges. XAI can uncover hidden patterns in the data that contribute to inconsistencies or ambiguities, providing insights into how ontologies can be refined and validated. For example, techniques such as SHAP (SHapley Additive exPlanations; Lundberg & Lee [16]) and LIME (Local Interpretable Model-agnostic Explanations; Ribeiro et al. [17]) can elucidate the underlying drivers of classification decisions, shedding light on how constructs are delineated and organized within an ontology. Specifically, SHAP assigns importance values to features in a model, allowing researchers to understand which data points contribute most to the classification of a construct, whereas LIME provides local explanations for individual predictions, helping to identify inconsistencies or biases in the data. These techniques ensure that ontology learning remains interpretable and that the resulting ontologies are grounded in empirical evidence—a synergy that not only mitigates the limitations of ontology learning but also facilitates a more robust mapping of constructs to underlying data. For example, in refining the construct of impulsivity, SHAP can reveal which behavioral tasks or neural markers are most influential in defining the construct, while LIME can highlight context-specific nuances that may otherwise be overlooked.

Ontology learning implemented in this way can enhance the interpretability of both complex data and the AI models used to analyze them. Data-centric XAI focuses on extracting transparent, actionable insights from neural datasets, while model-centric XAI helps elucidate the inner workings of AI systems, fostering interdisciplinary trust and collaboration [18,19]. XAI can help ensure that models remain interpretable and that their outputs are meaningful in advancing neuroscience research, while facilitating collaboration between neuroscience and machine learning experts [20–22]. Recent efforts in neuroimaging, such as those aimed at creating large-scale cognitive architectures, highlight the potential of structured ontologies to organize complex datasets and refine our understanding of cognitive processes [11,15,23]. Frameworks built using ontology learning can harmonize methodologies across different approaches, such as task-based fMRI and resting-state connectivity studies, to uncover latent dimensions of cognition, like the individual components of executive function, and provide insight into their associated neural systems [24].

These frameworks are not only theoretically valuable but also offer practical benefits in clinical and applied neuroscience contexts; for example, research into the neural basis of cognitive control illustrates how structured approaches to construct definition can enhance therapeutic outcomes. Attention Deficit Hyperactivity Disorder (ADHD) interventions, which often focus on improving executive functions like inhibition and working memory, could be refined by identifying distinct cognitive processes associated with neural systems implicated in ADHD pathology, such as the fronto-striatal and fronto-parietal networks [25,26]. Such developments could support personalized therapeutic strategies, tailoring interventions to specific deficits identified through precise construct mappings.

The integration of ontology learning with other methodologies is also vital in advancing neuroscience toward a more systematic and unified science. Ontology learning relies on diverse data sources, including scientific literature, structured databases, and experimental datasets. Text mining tools such as natural language processing (NLP) algorithms play a critical role in identifying key terms, phrases, and their contexts from unstructured text. Neuroscience studies often use text mining to extract cognitive terms from articles, identify co-occurrence patterns, and infer relationships between constructs like cognitive control and response inhibition [14]. Structured data repositories offer an alternative source for ontology development, systematically organizing constructs, tasks, and their neural correlates to provide a foundation for formal ontology creation.

4. Current uses of ontologies in psychology and neuroscience

The integration of ontology learning into psychological research helps enhance theoretical clarity and methodological rigor by formalizing psychological constructs and their relationships. In particular, it can address critical issues such as task-construct conflation and inconsistent terminology across subfields. For example, working memory is often measured using tasks like the N-back or complex digit span, which may engage distinct neural processes [27]. Ontologies help disentangle these relationships, linking constructs to their neural substrates systematically [11], enabling more precise hypotheses about brain-behavior relationships and advancing experimental rigor.

Relatedly, constructs like executive function or cognitive control are often operationalized differently across studies, leading to challenges in synthesis and interpretation [23,28]. These variations may stem from historical trends or subfield differences, complicating the identification of similarities across labels [29,30]. Ontology learning systematically evaluates these operationalizations by mapping them to a standardized framework, revealing inconsistencies and redundancies. For instance, tasks like the Stroop and Wisconsin Card Sorting Test are often used interchangeably to measure cognitive control and flexibility, yet they involve distinct processes [31]. Ontology learning provides tools to disentangle these overlaps, guiding the development of tasks that isolate individual constructs more precisely. AI-driven models, such as graph-based reasoning algorithms, have been employed in bioinformatics in this way to link experimental results to theoretical frameworks [4], and similar techniques in psychology could enhance hypothesis generation and testing.

Furthermore, efforts like the Cognitive Atlas demonstrate the utility of ontological approaches in clarifying relationships between constructs and tasks [11]. This open-science platform provides a structured vocabulary for cognitive terms, tasks, mental processes, and disorders, standardizing definitions of constructs like working memory and executive control, which vary across studies [11]. Linking constructs to specific tasks, such as the N-Back test for working memory, has facilitated cumulative research and improved reproducibility in cognitive neuroscience [32,33], and integration with tools like NeuroVault (neurovault.org) has enabled mapping cognitive constructs onto brain imaging data, bridging psychological and neural domains [34]. Beyond task design, ontology learning can assist in integrating neural and behavioral data, a critical goal in cognitive neuroscience. Data from the Human Connectome Project has helped document the associated cognitive constructs and neural connectivity patterns [35], while tools like NeuroSynth (neurosynth.org; Yarkoni et al. [36,37]) and BrainMap (brainmap.org; Vanasse et al. [38]) leverage ontological frameworks to map constructs to brain activity, refining these mappings further.

Ontology learning also has significant translational implications, particularly in understanding and treating psychological disorders. Via its ability to formalize constructs like depression or anxiety, ontologies disentangle heterogeneous symptom profiles, identifying subtypes based on symptom clustering and neural correlates [39]. This stratification links subtypes to genetic predispositions, environmental influences, and treatment responses [40], ultimately enabling personalized treatment strategies. Ontologies can also seamlessly integrate multi-modal data, such as neuroimaging and genetic markers, to enhance subtype classification precision [15]. These frameworks foster cross-disciplinary collaboration, providing a shared vocabulary for advancing diagnostic criteria and therapeutic outcomes. As such, ontology learning not only enhances terminological precision but also lays the groundwork for interdisciplinary collaboration, enabling neuroscience to align more closely with fields that require structured and interpretable knowledge bases [21,22].

5. Interdisciplinary implications of ontology learning

Beyond psychology and neuroscience, ontology learning also helps bridge the gap with other disciplines, fostering collaboration and innovation. It is deeply intertwined with AI, particularly in knowledge representation and semantic reasoning. Techniques like natural language processing (NLP), machine learning, and automated reasoning underpin ontology construction [41,42]. In turn, psychological constructs mapped through ontology learning inform AI systems, enhancing their ability to model cognition and decision-making [43]. This bidirectional exchange refines both AI systems and psychological theories.

Interdisciplinary collaboration can be facilitated through the establishment of shared platforms and frameworks that enable researchers from different fields to contribute to ontology development. For instance, the Open Biological and Biomedical Ontology (OBO) Foundry provides a model for collaborative ontology development, bringing together experts from biology, medicine, and computer science [44]. These collaborations not only enhance psychological research but also facilitate knowledge translation into education, human-computer interaction, and public health. Their interdisciplinary nature fosters methodological and theoretical innovations, addressing challenges that single disciplines cannot resolve alone.

Neuroinformatics relies on ontological frameworks to organize complex datasets from neuroimaging and electrophysiological studies. Projects like the Human Brain Project [45] and Allen Brain Atlas [46,47] use ontologies to map data onto standardized brain regions and networks. Ontology learning can further integrate neuroinformatics with bioinformatics, uncovering genetic markers linked to cognitive and emotional traits [4]. This supports personalized medicine by identifying genetic predispositions for mental health conditions and tailoring interventions accordingly. The Human Connectome Project (HCP) exemplifies this, using ontology learning to annotate data with constructs like attention and self-regulation, linking them to neural circuits [38]. For instance, one study validated the distinctiveness of executive functions by connecting task performance to network connectivity patterns [48].

Ontology learning also allows exploration of the relationship between language and cognition, mapping linguistic structures onto psychological constructs. Research on semantic memory and language processing has used ontologies to model how concepts are represented in the brain [11,49,50]. This has practical implications, such as improving language-based interventions for neurological disorders like aphasia or Alzheimer's disease [51]. In addition, ontology learning intersects with philosophy and ethics, with the potential to clarify debates about constructs like intelligence or mental illness by formalizing their definitions [52]. It also informs ethical guidelines for AI applications in neuroscience, ensuring constructs are accurately defined to reduce misuse [53].

In translational research, ontology learning aids in developing diagnostic tools and interventions. The Research Domain Criteria (RDoC) framework, for example, redefines mental health disorders based on observable behavior and neurobiological measures [54]. Ontology learning has enabled nuanced classifications of disorders like depression and schizophrenia, identifying subtypes with distinct neural and symptom profiles for targeted treatments [55]. This approach standardizes mental health constructs, facilitating communication across neuroscience, psychology, and pharmacology [56]. Ontology-based tools have already harmonized clinical terminologies like ICD and DSM [57], improving diagnostic accuracy and treatment efficacy.

Finally, ontology learning aligns with open science principles, enhancing data sharing and reproducibility. Via standardized frameworks, ontologies facilitate collaboration across disciplines [33]: initiatives like the Biomedical Informatics Research Network (BIRN) have used ontologies to harmonize metadata, enabling replication and validation of findings [58]. Expanding such efforts through ontology learning can further enhance transparency and interdisciplinary collaboration.

6. Limitations and challenges

While ontology learning holds significant promise for refining psychological constructs and advancing neuroscience, several challenges must be addressed before widespread adoption. First, psychological constructs are inherently multi-dimensional and lack universally accepted definitions, leading to variability in their operationalization across studies. For example, cognitive control is measured using diverse tasks like the Stroop, Go/No-Go, and Flanker tasks, each tapping into overlapping but distinct processes [11,31]. This variability complicates the creation of clear, non-redundant ontologies [23]. XAI can help by offering transparency into how algorithms parse constructs, identifying overlaps and gaps. In addition, constructs are often polysemous, with meanings varying by context (e.g., attention in cognitive vs. clinical studies), increasing misclassification risks [12]. XAI frameworks like SHAP [16] can clarify how models weigh task data, aiding in refining operational definitions. A related challenge is disentangling tasks from constructs. Many neuroscience tasks, such as the aforementioned N-back task, conflate multiple processes (e.g., working memory, attention, and executive control), complicating ontology integration [13,27]. Moreover, data quality varies, with noisy and hard-to-replicate results in psychology posing challenges for reliable ontology construction [59].

Psychological constructs are also subjective and context-dependent, making consistent ontological relationships difficult to establish. For instance, constructs like personality and mental health are influenced by social and cultural factors, leading to variations across populations and time [60]. The concept of mental disorder, for example, varies across cultural contexts and diagnostic frameworks (e.g., DSM-5 vs. ICD-10). Ontology learning tools must account for this diversity, but flexibility can lead to generalization issues. In addition, constructs evolve with new empirical findings, requiring iterative ontology updates that may introduce inconsistencies.

Integrating heterogeneous data sources, such as behavioral data, neuroimaging, genetics, and self-reports, poses another challenge. Different data modalities vary in granularity and may not map cleanly onto constructs like decision-making, which are influenced by contextual factors [61]. Current AI methods, including NLP and machine learning, struggle to merge these data types meaningfully, and small sample sizes in neuroscience further complicate robust model training. In this context, techniques such as data augmentation and transfer learning can greatly enhance the robustness of ontology learning [62]. Ethical concerns can also arise, particularly with sensitive data like neuroimaging or genetic information, as privacy and consent issues are critical, especially for clinical or vulnerable populations. Biases in data or algorithms can skew representations, perpetuating stereotypes or inequalities [63]; for example, neuroimaging data predominantly comes from Western, educated, industrialized, rich, and democratic (WEIRD) populations, which can limit generalizability [64]. Relatedly, another challenge lies in the potential biases inherent in selecting data inputs or sources for ontology learning: pre-defined data inputs may encode biases or incomplete representations of the domain of interest [65]. To mitigate these risks, the field should develop governance frameworks that ensure transparency, accountability, and fairness in ontology learning. These frameworks could include guidelines for data anonymization, informed consent, and bias detection. XAI can also play a key role in identifying and mitigating biases in datasets, ensuring that ontology learning systems remain fair and generalizable [62].

Finally, technical limitations, such as the lack of transparency in AI models, also hinder ontology learning. The black-box nature of many machine learning models undermines trust, especially in clinical psychology [66]. XAI can address this by providing interpretable explanations, such as saliency maps, but computational resource requirements remain a barrier, particularly in underfunded fields [67]. To ensure transparency in XAI-driven ontology learning, we recommend the adoption of standardized reporting guidelines that detail the methodologies used, the data sources involved, and the interpretability techniques applied. These guidelines should be developed in collaboration with the broader scientific community and should be made publicly available to facilitate reproducibility and trust in the resulting ontologies.

Moreover, scalability and maintenance are long-term challenges. Ontologies must evolve with new research, but dynamic updates can lead to inconsistencies across versions (Sure et al., 2009). XAI can address this challenge by automating parts of the ontology construction process, reducing the need for extensive human oversight. For instance, machine learning algorithms can be used to pre-process large datasets, identify patterns, and generate preliminary ontologies, which can then be refined by domain experts. This hybrid approach not only enhances scalability, but also ensures that the resulting ontologies are both accurate and interpretable. In this changing landscape, maintaining ontology integrity will require ongoing collaboration among researchers, domain experts, and data scientists.

7. Future directions

The integration of ontology learning into refining psychological constructs and advancing neuroscience offers exciting opportunities for future research. As AI and ontology learning evolve, progress in several key areas can enhance our understanding of complex psychological phenomena and improve the precision of psychological constructs in neuroscience.

First, the continued development of machine learning, particularly deep learning, holds significant promise for ontology learning. Deep learning algorithms excel at modeling complex, non-linear relationships, making them well-suited for uncovering hidden patterns in large datasets, such as those from neuroimaging or behavioral research [43]. As AI algorithms advance, they could help generate more accurate and nuanced ontological representations of constructs like cognition, emotion, and mental health [68]. Hybrid models combining rule-based systems with machine learning could also improve ontology interpretability and transparency [69,70]. Advances in XAI will further enable the creation of interpretable models tailored to neuroscience's complexities: quantitative metrics can serve to evaluate the effectiveness of ontology development in terms of interpretability (e.g., SHAP values) and consistency (e.g., alignment with existing ontologies), while qualitative metrics could involve expert reviews to assess the accuracy and relevance of the updated ontologies. Relatedly, cross-disciplinary collaboration will be essential for the future of ontology learning. Successful ontology development requires input from psychology, neuroscience, data science, and philosophy [71], with open science initiatives and data-sharing platforms, playing a key role in enhancing the generalizability and precision of ontologies by identifying patterns across studies and populations.

Furthermore, future ontology learning will likely focus on integrating multi-modal data, which is increasingly common in psychology and neuroscience. Combining neuroimaging, genetic, behavioral, and real-time data from wearable devices requires ontologies to represent these diverse data types cohesively [11]. Multi-modal data integration enables a more holistic understanding of psychological constructs, capturing their dynamic, multidimensional nature [72]. In this context, ethical considerations will become increasingly important as ontology learning is applied in psychology and neuroscience. Rigorous ethical frameworks are needed to address concerns such as data privacy, bias, and misuse [63]: developing fair and transparent AI systems, including algorithms that detect and mitigate bias, will be crucial for ensuring responsible use and preventing the perpetuation of stereotypes. Fostering educational initiatives to train researchers in ontology learning and XAI interpretation will therefore be essential for the widespread adoption and effective application of these methodologies.

Finally, dynamic, real-time ontologies represent a promising direction. As wearable devices and smartphones collect real-time psychological and physiological data, ontologies can evolve dynamically to reflect new findings [73]. This approach is particularly valuable for studying mental health, where rapid changes in psychological states can require continuous monitoring. Dynamic ontologies also support personalized interventions, tailoring treatments to individuals' unique profiles [74]. To be accurate and impactful, future ontology learning will need to incorporate cross-cultural considerations; as constructs can vary substantially across cultures [75]—incorporating contextual information into ontology learning will be essential for ensuring that ontologies are equitable and globally applicable.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: David Moreau reports was provided by The University of Auckland. David Moreau reports a relationship with The University of Auckland that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J.W. Dalley, B.J. Everitt, T.W. Robbins, Impulsivity, compulsivity, and top-down cognitive control, Neuron 69 (4) (2011) 680–694.
[2] D.M. Dick, G. Smith, P. Olausson, S.H. Mitchell, R.F. Leeman, S.S. O'Malley, K. Sher, Understanding the construct of impulsivity and its relationship to alcohol use disorders, Addict. Biol. 15 (2) (2010) 217–226.
[3] L. Sharma, K.E. Markon, L.A. Clark, Toward a theory of distinct types of impulsive behaviors: a meta-analysis of self-report and behavioral measures, Psychol. Bull. 140 (2) (2014) 374–408.
[4] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology, The Gene Ontology Consortium, Nat. Genet. 25 (1) (2000) 25–29.
[5] R. Cornet, N. de Keizer, Forty years of SNOMED: a literature review, BMC Med. Inform. Decis. Mak. 8 (2008) 1–6.
[6] D. Lee, N. de Keizer, F. Lau, R. Cornet, Literature review of SNOMED CT use, J. Am. Med. Inform. Assoc. 21 (e1) (2013) e11–e19.
[7] M. Bota, H.W. Dong, L.W. Swanson, Brain architecture management system, Neuroinformatics 3 (1) (2005) 15–48.
[8] T.R. Gruber, A translation approach to portable ontologies, Knowl. Acquis. 5 (2) (1993) 199–220.
[9] J.B. Bard, S.Y. Rhee, Ontologies in biology: design, applications, and future challenges, Nat. Rev. Genet. 5 (3) (2004) 213–222.
[10] P. Cimiano, A. Hotho, S. Staab, Learning concept hierarchies from text corpora using formal concept analysis, J. Artif. Intell. Res. 24 (1) (2005) 305–339.
[11] R.A. Poldrack, A. Kittur, D. Kalar, E. Miller, C. Seppa, Y. Gil, D.S. Parker, F.W. Sabb, R.M. Bilder, The cognitive atlas: toward a knowledge foundation for cognitive neuroscience, Front. Neuroinform. 5 (2011) 17.
[12] T. Yarkoni, J. Westfall, Choosing prediction over explanation in psychology: lessons from machine learning, Perspect. Psychol. Sci. 12 (6) (2016) 1100–1122.
[13] I.W. Eisenberg, P.G. Bissett, A. Zeynep Enkavi, J. Li, D.P. MacKinnon, L.A. Marsch, R.A. Poldrack, Uncovering the structure of self-regulation through data-driven ontology discovery, Nat. Commun. 9 (1) (2019) 2319.
[14] J.A. Turner, A.R. Laird, The cognitive paradigm ontology: design and application, Neuroinformatics 10 (1) (2012) 57–66.
[15] C.W. Woo, L.J. Chang, M.A. Lindquist, T.D. Wager, Building better biomarkers: brain models in translational neuroimaging, Nat. Neurosci. 20 (3) (2017) 365–377.
[16] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017), 2017, pp. 4765–4774.
[17] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?" Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.
[18] F. Doshi-Velez, B. Kim, Towards a rigorous science of interpretable machine learning, arXiv preprint arXiv:1702.08608, 2017.
[19] C. Rudin, Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead, Nat. Mach. Intell. 1 (5) (2019) 206–215.
[20] D. Gunning, D. Aha, DARPA's explainable artificial intelligence (XAI) program, AI Mag. 40 (2) (2019) 44–58.
[21] W. Samek, T. Wiegand, K.R. Müller, Explainable artificial intelligence: understanding, visualizing, and interpreting deep learning models, arXiv:1708.08296, 2017.
[22] E. Tjoa, C. Guan, A survey on explainable artificial intelligence (XAI): towards medical XAI, IEEE Trans. Neural Netw. Learn. Syst. 32 (11) (2020) 4793–4813.
[23] D. Moreau, K. Wiebels, Psychological constructs as local optima, Nat. Rev. Psychol. 1 (2022) 188–189.
[24] D.E. Nee, S. Kastner, J.W. Brown, Functional heterogeneity of conflict, error, task-switching, and unexpectedness effects within medial prefrontal cortex, NeuroImage 54 (1) (2011) 528–540.
[25] R.A. Barkley, Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD, Psychol. Bull. 121 (1) (1997) 65–94.
[26] D.A. Fair, D. Bathula, M.A. Nikolas, J.T. Nigg, Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD, Proc. Natl. Acad. Sci. USA 109 (17) (2012) 6769–6774.
[27] M.J. Kane, A.R.A. Conway, T.K. Miura, G.J.H. Colflesh, Working memory, attention control, and the *n*-back task: a question of construct validity, J. Exp. Psychol. Learn. Mem. Cogn. 33 (3) (2007) 615–622.
[28] E.I. Fried, Moving forward: how depression heterogeneity hinders progress in treatment and research, Expert Rev. Neurother. 17 (5) (2017) 423–425.
[29] D. Moreau, Shifting minds: a quantitative reappraisal of cognitive intervention research, Perspect. Psychol. Sci. 16 (1) (2021) 148–160.
[30] D. Moreau, How malleable are cognitive abilities? A critical perspective on popular brief interventions, Am. Psychol. 77 (3) (2022) 409–423.
[31] F.W. Sabb, C.E. Bearden, D.C. Glahn, D.S. Parker, N.B. Freimer, R.M. Bilder, A collaborative knowledge base for cognitive phenomics, Mol. Psychiatry 13 (4) (2008) 350–360.
[32] D. Moreau, B. Gamble, Conducting a meta-analysis in the age of open science: tools, tips, and practical recommendations, Psychol. Methods 27 (3) (2022) 426–432.
[33] R.A. Poldrack, et al., Scanning the horizon: towards transparent and reproducible neuroimaging research, Nat. Rev. Neurosci. 18 (2) (2017) 115–126.
[34] K.J. Gorgolewski, G. Varoquaux, G. Rivera, Y. Schwartz, V.V. Sochat, S.S. Ghosh, C. Maumet, T.E. Nichols, J.B. Poline, T. Yarkoni, D.S. Margulies, R.A. Poldrack, NeuroVault.org: a repository for sharing unthresholded statistical maps, parcellations, and atlases of the human brain, NeuroImage 124 (2016) 1242–1244.
[35] D.C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T.E. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S.W. Curtiss, S. Della Penna, D. Feinberg, M.F. Glasser, N. Harel, A.C. Heath, L. Larson-Prior, D. Marcus, G. Michalareas, S. Moeller, et al., The human connectome project: a data acquisition perspective, NeuroImage 62 (4) (2012) 2222–2231.
[36] T. Yarkoni, R.A. Poldrack, T.E. Nichols, D.C. Van Essen, T.D. Wager, Large-scale automated synthesis of human functional neuroimaging data, Nat. Methods 8 (8) (2011) 665–670.
[37] T. Yarkoni, R. Poldrack, T. Nichols, D. Van Essen, T. Wager, NeuroSynth: a new platform for large-scale automated synthesis of human functional neuroimaging data, in: Conference Abstract: 4th INCF Congress of Neuroinformatics, Front. Neuroinform. (2011).
[38] T.J. Vanasse, P.M. Fox, D.S. Barron, M. Robertson, S.B. Eickhoff, J.L. Lancaster, P.T. Fox, BrainMap VBM: an environment for structural meta-analysis, Hum. Brain Mapp. 39 (8) (2018) 3308–3325.
[39] K.S. Kendler, J. Myers, The genetic and environmental relationship between major depression and the five-factor model of personality, Psychol. Med. 40 (5) (2010) 801–806.
[40] A.T. Drysdale, L. Grosenick, J. Downar, K. Dunlop, F. Mansouri, Y. Meng, R.N. Fetcho, B. Zebley, D.J. Oathes, A. Etkin, A.F. Schatzberg, K. Sudheimer, J. Keller, H.S. Mayberg, F.M. Gunning, G.S. Alexopoulos, M.D. Fox, A. Pascual-Leone, H.U. Voss, et al., Resting-state connectivity biomarkers define neurophysiological subtypes of depression, Nat. Med. 23 (1) (2017) 28–38.
[41] T. Mikolov, et al., Efficient estimation of word representations in vector space, arXiv:1301.3781, 2013.
[42] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, 2019, pp. 4171–4186.
[43] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[44] B. Smith, et al., The OBO foundry: coordinated evolution of ontologies to support biomedical data integration, Nat. Biotechnol. 25 (11) (2007) 1251–1255.
[45] K. Amunts, C. Ebell, J. Muller, M. Telefont, A. Knoll, T. Lippert, The human brain project: creating a European research infrastructure to decode the human brain, Neuron 92 (3) (2016) 574–581.
[46] E.S. Lein, et al., Genome-wide atlas of gene expression in the adult mouse brain, Nature 445 (2007) 168–176.
[47] S.W. Oh, et al., A mesoscale connectome of the mouse brain, Nature 508 (2014) 207–214.
[48] M.W. Cole, J.R. Reynolds, J.D. Power, G. Repovs, A. Anticevic, T.S. Braver, Multitask connectivity reveals flexible hubs for adaptive task control, Nat. Neurosci. 16 (11) (2013) 1348–1355.
[49] R.A. Poldrack, T. Yarkoni, From brain maps to cognitive ontologies: informatics and the search for mental structure, Annu. Rev. Psychol. 67 (1) (2016) 587–612.
[50] M.A.L. Ralph, E. Jefferies, K. Patterson, C. Tompkins, The neural and computational bases of semantic cognition, Nat. Rev. Neurosci. 18 (2017) 42–55.
[51] M. O'Sullivan, S. Brownsett, D. Copland, Language and language disorders: neuroscience to clinical practice, Pract. Neurol. 19 (5) (2019) 380–388.
[52] R. Cooper, Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders, Springer, 2005.
[53] L. Floridi, et al., AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations, Minds Mach. 28 (4) (2018) 689–707.
[54] T. Insel, B. Cuthbert, M. Garvey, R. Heinssen, D.S. Pine, K. Quinn, C. Sanislow, P. Wang, Research domain criteria (RDoC): toward a new classification framework for research on mental disorders, Am. J. Psychiatr. 167 (7) (2010) 748–751.
[55] B.N. Cuthbert, The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology, World Psychiatry 13 (1) (2014) 28–35.
[56] T. Insel, B. Cuthbert, Brain disorders? Precisely, Science 348 (6234) (2015) 499–500.
[57] J.J. Cimino, Desiderata for controlled medical vocabularies in the twenty-first century, Methods Inf. Med. 37 (4–5) (1998) 394–403.
[58] D.B. Keator, et al., A national human neuroimaging collaboratory enabled by the biomedical informatics research network (BIRN), IEEE Trans. Inf. Technol. Biomed. 12 (2) (2008) 162–172.

[59] H. Pashler, E. Wagenmakers, Editors' introduction to the special section on replicability in psychological science: a crisis of confidence?, Perspect. Psychol. Sci. 7 (6) (2012) 528–530.
[60] D.C. Funder, Personality, Annu. Rev. Psychol. 52 (2001) 197–221.
[61] L. Pessoa, How do emotion and motivation direct executive control?, Trends Cogn. Sci. 13 (4) (2009) 160–166.
[62] S. Barocas, M. Hardt, A. Narayanan, Fairness and machine learning, FairMLBook. Retrieved from https://fairmlbook.org/, 2019.
[63] C. O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy, Crown Publishing Group, 2016.
[64] J. Henrich, S.J. Heine, A. Norenzayan, The weirdest people in the world?, Behav. Brain Sci. 33 (2–3) (2010) 61–83, discussion 83–135.
[65] C. Brewster, F. Ciravegna, Y. Wilks, Knowledge acquisition for ontology construction, in: S. Handschuh, S. Staab (Eds.), Annotation for the Semantic Web, IOS Press, 2004, pp. 155–172.
[66] Z.C. Lipton, The mythos of model interpretability, in: 2016 ICML Workshop on Human Interpretability in Machine Learning, 2016, New York, NY.
[67] Y. Bengio, Learning deep architectures for AI, Found. Trends Mach. Learn. 2 (1) (2015) 1–127.
[68] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, J. Dean, A guide to deep learning in healthcare, Nat. Med. 25 (1) (2019) 24–29.
[69] L. Jiajia, X. Han, Y. Qin, F. Tan, Y. Chen, Z. Wang, H. Song, X. Zhou, Y. Zhang, L. Hu, P. Hu, Artificial intelligence accelerates multi-modal biomedical process: a survey, Neurocomputing 558 (2023) 126720.
[70] A.C. Khadir, H. Aliane, A. Guessoum, Ontology learning: grand tour and challenges, Comput. Sci. Rev. 39 (2021) 100339.
[71] K. Okamura, Interdisciplinarity revisited: evidence for research impact and dynamism, Palgrave Commun. 5 (2019) 141.
[72] A. Kline, H. Wang, Y. Li, Multimodal machine learning in precision health: a scoping review, npj Digit. Med. 5 (2022) 171.
[73] J. Zulueta, A.D. Leow, O. Ajilore, Real-time monitoring: a key element in personalized health and precision health, Focus 18 (2) (2020) 175–180.
[74] C. Weyant, M.L. Brandeau, Personalization of medical treatment decisions: simplifying complex models while maintaining patient health outcomes, Med. Decis. Mak. 42 (4) (2022) 450–460.
[75] S.J. Heine, Cultural Psychology, W.W. Norton & Company, 2016.