nature reviews neuroscience

https://doi.org/10.1038/s41583-023-00778-7

Review article

Check for updates

Causation in neuroscience: keeping mechanism meaningful

Lauren N. Ross [1] ✉ & Dani S. Bassett [2,3,4,5,6,7]

Abstract

A fundamental goal of research in neuroscience is to uncover the causal structure of the brain. This focus on causation makes sense, because causal information can provide explanations of brain function and identify reliable targets with which to understand cognitive function and prevent or change neurological conditions and psychiatric disorders. In this research, one of the most frequently used causal concepts is 'mechanism' — this is seen in the literature and language of the field, in grant and funding inquiries that specify what research is supported, and in journal guidelines on which contributions are considered for publication. In these contexts, mechanisms are commonly tied to expressions of the main aims of the field and cited as the 'fundamental', 'foundational' and/or 'basic' unit for understanding the brain. Despite its common usage and perceived importance, mechanism is used in different ways that are rarely distinguished. Given that this concept is defined in different ways throughout the field — and that there is often no clarification of which definition is intended — there remains a marked ambiguity about the fundamental goals, orientation and principles of the field. Here we provide an overview of causation and mechanism from the perspectives of neuroscience and philosophy of science, in order to address these challenges.

Sections

Introduction

Why do causation and mechanism matter for neuroscience?

How is 'mechanism' used in neuroscience?

What are the challenges and implications of this variable usage?

How can we address these challenges?

Causation in neuroscience: mechanism and its alternatives

[1]Department of Logic and Philosophy of Science, University of California, Irvine, Irvine, CA, USA. [2]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA. [3]Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA. [4]Department of Physics & Astronomy, University of Pennsylvania, Philadelphia, PA, USA. [5]Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA. [6]Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA. [7]Santa Fe Institute, Santa Fe, NM, USA. ✉e-mail: rossl@uci.edu

Introduction

A central aim of neuroscientific research is to clarify the causal structure of the brain, be that at the lower scales of molecular and cellular interactions or the higher scales of neural circuitry, brain regions and macro-scale networks. This interest in causation is understandable because of the invaluable role of causes in supporting explanations of brain dynamics and associated behavioural outcomes. As causes provide control over their effects, they identify reliable targets to explain, understand, change and prevent brain-related outcomes of interest[1–5]. Identifying different types of causal systems, information and patterns in the brain — and having a robust theoretical framework for capturing their features and implications — is vital for progress in neuroscience and its sibling disciplines of neurology and psychiatry.

One of the most frequently used causal concepts in neuroscience is 'mechanism'. This concept is found in the literature and language of the field, in grant and funding inquiries that specify what research is supported, and in journal guidelines detailing which contributions are considered for publication (Tables 1 and 2). In these contexts, it is often suggested that a deep explanation and understanding of the brain is provided by identifying its basic and fundamental mechanisms[6,7]. In current neuroscientific research, examples of the mechanism concept include "cellular mechanisms", "circuit mechanisms", "systems-level mechanisms", "cognitive mechanisms" and "network mechanisms", to name a few[3,8–19]. More concrete illustrations are the "ionic mechanisms of the action potential", "mechanisms of signal propagation", "transcriptional mechanisms", "mechanisms of neuromodulation", "barrier mechanisms", "mechanisms of cognitive processing" and reference to various "circuit mechanisms" and "network mechanisms" of the brain[3,8–13,20–23]. The term 'mechanism' not only has scientific appeal but also receives enormous attention in the theory and philosophy of neuroscience[1,24–26]. In these areas, it is common to find claims that genuine explanations in neuroscience always require the elucidation of mechanistic information about the brain, where mechanistic information is understood as lower-scale causal detail that produces the brain outcome of interest[24,25].

However, although the mechanism concept is exceedingly common in neuroscience, it is used in completely different ways that are rarely specified or distinguished. This myriad usage results in miscommunication, unintended meanings and a lack of conceptual clarity in the field. Such confusions can have dire implications when this term is used in publication regulations, dissemination of research findings, exploratory studies of complex causal systems in the brain, and funding solicitations that inspire and motivate new research. Given that the field lacks a clear, shared meaning of such a fundamental term, these current issues are inevitable and have numerous, field-wide implications.

In this Review, we provide an overview of causation and mechanism from the perspectives of neuroscience and philosophy of science. We begin by reviewing the philosophy of science literature on causation, explanation and the concept of mechanism. We then describe standard uses of this concept in neuroscience, including their distinct assumptions about the type, amount and scale of causal detail required to provide understanding. After presenting different meanings of mechanism, we examine key challenges that this variable usage produces, whether they occur in basic research, funding priorities or communication to expert and public audiences. Next, we draw on the scientific, theoretical and philosophical points made to provide suggestions for moving forward. We close by discussing the need for clarity about causation to capture the main principles in the field and to promote advances in our understanding of the causal structure of the brain.

Why do causation and mechanism matter for neuroscience?

Causation has been a long-standing topic of interest in science, philosophy and many other domains. A major reason for this focus is that causal relationships provide explanations and understanding of 'how' a system — such as the brain, a neural ensemble or a cellular circuit — works, functions and behaves. This provision arises because causes explain their effects by clarifying the inner workings and steps that lead to these outcomes. Instead of merely describing, predicting or classifying neural and brain outcomes, causal models identify factors that control, explain and 'make a difference' to them[1,27]. These causal factors support deep understanding and explanation, and guide our efforts to change and control outcomes, which are central in our interests to prevent and treat brain disorders and diseases. The classification of distinct entities in neuroscience is common and serves various goals, but such classifications do not require causal information. Furthermore, consider descriptive models, which include a physical instantiation of the system, and phenomenological models, which capture conceptual components of a system (without requiring physical instantiation). Mere descriptive and phenomenological models can represent and re-describe the outcome of interest, but without causal information they fall short of explaining such outcomes or of offering the knowledge necessary to change or prevent outcomes[6,7]. In other words, models that are merely descriptive or phenomenological are often "rather limited" as they can do "a good job of characterizing what the system does, but they have no idea how and why"[7,28].

The importance of causation raises many questions that remain vigorously debated in the philosophy of neuroscience. These include questions about how to define causation, what methods successfully identify or establish it, and what scale and type of causal information is needed to understand how neural systems work[29–34]. A philosophy of science approach to these questions aims to specify the principled methods, concepts and assumptions that scientists use to gain understanding of the world. This effort involves a descriptive aim of capturing concepts and methods in science, but also a normative aim that provides guidance for how these concepts and methods should (and should not) be understood and used.

Discussions of causal explanation in the philosophy of neuroscience place considerable attention on the mechanism concept. The modern use of mechanism originates in the seventeenth-century mechanist philosophy views of Descartes, Newton, Boyle and others[35]. These mechanist programmes provided an alternative to vitalist conceptions of living systems, which cited vital forces, occult powers and magical properties in accounting for natural phenomena[36]. By contrast, the mechanist framework explained living systems with "the mathematical discipline of mechanics" in which systems consist of small entities that mechanically interact through "contact action" to produce all system-level outcomes[35]. This framework was reductive in citing causes at lower scales and in confining conceptions of causation to action, physical forces and matter in motion. Living systems were analogized to machines — viewed as bête machines (beast machines) — comprising lower-scale corpuscles that mechanically produce behaviours, similar to levers, pulleys and pipes in simple machines[35,37]. This seventeenth-century mechanist framework has substantially influenced the modern life sciences. This influence is evidenced by common appeal to mechanisms, frequent analogies to machines and mechanical systems, and the prevalence of reductive

Table 1 | Examples of publication guidelines from neuroscience and related journals that mention the 'mechanism' concept

Journal	Publication guidelines	Ref.
Annals of Neurology	"The *Annals of Neurology* publishes articles of broad interest in neurology, particularly those with high impact in understanding the mechanisms and treatment of diseases of the human nervous system." "*Annals* rarely publishes case reports, unless they definitively identify a new mechanism of disease or treatment."	133
Biological Psychiatry	"Reviews should be novel and have sufficient supporting literature, which should be integrated into a mechanistic model when applicable."	134
Experimental Neurology	"*Experimental Neurology*, a Journal of Neuroscience Research, publishes original research in neuroscience with a particular emphasis on novel findings in neural development, regeneration, plasticity and transplantation. The journal has focused on research concerning basic mechanisms underlying neurological disorders." "Research Articles — the Journal publishes Research Articles focusing on experimental models of neurological and psychiatric diseases. These should describe significant, new and carefully confirmed findings with attention to mechanisms of diseases and/or treatments."	135
The Lancet Neurology	"We invite submission of all clinical trials, whether phase 1, 2, 3, or 4 (see *Lancet* 2006; 368: 827–28). For phase 1 trials, we especially encourage those of a novel substance for a novel indication, if there is a strong or unexpected beneficial or adverse response, or a novel mechanism of action."	136

Passages of text quoted with permission from ref. 133, Wiley;[134–136], Elsevier.

explanation, in which explanation requires the citation of causes at lower scales.

The mechanism concept continues to flourish in modern neuroscience, and philosophers have taken note[24,38–42]. About two decades ago, growing philosophical views introduced a 'new' mechanistic account, which claims that genuine causal explanations in neuroscience must cite mechanisms[24,25]. This account is considered 'new' to distinguish it from the seventeenth-century mechanists, but it builds on their work in articulating a modern notion of mechanism[43]. Early work in this area claimed that mechanisms are causal systems with specific features. Similar to the original seventeenth-century notion of mechanism, many suggested that mechanisms are hierarchical systems with lower-scale causal parts, that they contain substantial causal detail and that they involve mechanical causal relationships, in terms of force, action and motion[44]. This notion also captured common analogies of neurological systems to machines, as many machines in everyday life have these three main features. Other notions of mechanism are more aetiological, highlighting causal intermediates that lie between a cause and its effect[43,45]. Both hierarchical and aetiological notions of mechanism resist counting abstract, macro-scale causal systems as genuine mechanisms. The reason is that genuine causal mechanisms are expected to contain considerable amounts of causal information and often from lower scales. Of course, these projects imply (often explicitly) that causal explanations in neuroscience require these types of lower-scale, fine-grained causal detail.

As work in this area has continued, many have suggested that there are limits to the mechanism concept and that there are other distinct types of causal system in neuroscience that are important[26,44,46–53]. In the philosophical literature, criticisms of the mechanism framework have argued against views that all explanations are mechanistic and against the view that "living systems can be fully explained by describing the causal mechanisms that are said to be operating within them"[29,54]. Some of this critical work has emphasized "process" and "processual" views of living systems, which capture causal systems that change, are more dynamic than fixed and lack the "rigid connections" assumed in the mechanistic perspective[54,55]. Process accounts identify issues with the assumption that living systems are well understood with a machine-like conception of mechanism[47]. Other philosophical criticisms of the mechanistic account suggest that this work should appreciate and be open to other causal systems in neuroscience, such as pathways, circuits and causal topologies[26,44,52]. In responding to these criticisms, mechanists have argued that the notion of mechanism should simply be broadened to encompass a range of other types of causal systems[43,56,57]. Instead of mechanism capturing a subset or class of causal systems, some scholars use it as a catch-all for any type of causal structure. This loosening of the mechanism concept is seen in current philosophical attempts to define mechanism, which equate mechanisms with nearly any type of causal system[43]. These philosophical attempts to broadly define mechanism have been criticized as in "serious danger of vacuity" and as guided by "imperialistic tendencies", as the term is used to encompass all causal systems in neuroscience, even when these causal systems appear to differ in important ways and have little in common[47,58]. The mechanism concept is found in many other scientific fields, including chemistry, biology, ecology, the social sciences, economics and the physical sciences[26,44,59–62]. In these domains it is often applied to complex, multi-component systems and is viewed as providing understanding about how these systems work and produce outcomes. Although similarities to other fields exist, gaining traction on mechanism in neuroscience often presents as more challenging, owing to the complex nature of the brain (amount of detail, distinct phenomena to capture, temporal and spatial scales, and so on), the extreme field-wide diversity of methods, tools, goals and topics of interest, and the importance of causal understanding to address brain health and disease.

A central topic to return to in this discussion is causation. Our explanations, understanding and notions of mechanism all depend on causality. Causation is the "raw material" out of which we get various causal structures and upon which our explanations and understanding depend[63]. The point is not to settle which definition of causation we should adopt but to keep this definition distinct from mechanism, which is reliant on causation. Although mechanistic explanation still receives attention in the literature, growing areas of work examine other non-mechanistic, causal systems in neuroscience and various forms of non-causal, mathematical explanation[26,44,47,64,65].

How is 'mechanism' used in neuroscience?

A rich variety of terms and concepts are used to refer to causal systems in neuroscience. These include mechanisms, pathways, cascades, circuits, networks, topologies and constraints, in addition to many others[6,66–79]. Among these, 'mechanism' is frequently used and often

Table 2 | Examples of grant and funding information that mentions the 'mechanism' concept and 'mechanistic research'

Grant or funding inquiries	Research funded	Ref.
National Institutes of Mental Health (NIMH), NIH	"NIMH requires an experimental therapeutics approach (see Clinical Trials – Applicant FAQs, Q3) for the development and testing of therapeutic, preventive, and services interventions, in which the studies evaluate not only the clinical effect of the intervention, but also generate information about the mechanisms underlying a disorder or an intervention response." "NIMH utilizes the NIH Parent R01 Clinical Trial Required, the NIH Parent R21 Clinical Trial Required, the NIH Parent R01 Basic Experimental Studies with Humans, and the NIH Parent R21 Basic Experimental Studies with Humans to accept only 'mechanistic' clinical trial applications (see below for descriptors of 'mechanistic' clinical trials)." "NIMH Support of Mechanistic Clinical Trials: NIH defines a mechanistic clinical trial as a study 'designed to understand a biological or behavioral process, the pathophysiology of a disease, or the mechanism of action of an intervention.'"	93
National Institute of Neurological Disorders and Stroke (NINDS), NIH	"The mission of the NINDS Division of Neuroscience (DON) is to support extramural research aimed at understanding fundamental mechanisms of development, structure and function of the nervous system in health and disease."	137
NSF Cognitive Neuroscience	"Supports research aimed at increasing understanding of the neural mechanisms of human cognition, including attention, learning, memory, decision making, language, social cognition and emotions." "The Cognitive Neuroscience Program seeks highly innovative proposals aimed at advancing a rigorous understanding of the neural mechanisms of human cognition."	138
NSF Collaborative Research in Computational Neuroscience (CRCNS)	"Supports collaborative research and data sharing that will advance the understanding of nervous system structure and function, mechanisms underlying nervous system disorders, and computational strategies used by the nervous system."	139
NSF Neural Systems Division of Integrative Organismal Systems Core Programs	"The Neural Systems Cluster supports mechanistic studies in neuroscience that span multiple levels of analysis ranging from the molecular and cellular to the complex behavioral aspects of organisms."	140

NIH, National Institutes of Health; NSF, National Science Foundation. Passages of text quoted with permission from ref. 93, NIMH;[137], NINDS,[138–140], NSF.

considered fundamental, basic and essential to our understanding of the brain[7]. Despite its common usage and perceived importance, the mechanism concept is used in different ways that are rarely distinguished. We identify three main uses of this term: a narrow (reductive) use, a broad (non-reductive) use and a non-causal use.

A first use of mechanism is narrow in that it refers to causal systems with particular features that are often reductive. These mechanisms "underlie", "underpin" or "implement" higher-scale systems and are characterized by microscale processes, physical–causal interactions, spatial-geometric features and an emphasis on fine-grained detail[6,7,80–83]. Concrete examples include ion channel mechanisms, receptor mechanisms, diffusion mechanisms and synaptic mechanisms of neural function, and more abstract references to mechanisms that are biophysical, biological and physiological (Fig. 1). Reference to this type of mechanism is most obvious in appeals to the 'underlying' mechanisms of some system — topology, network, model or physiological process — as the focus is on lower-scale causal factors that produce the behaviour of interest. These mechanisms emphasize the "biophysical" and "physical" causes that realize neural and brain systems — sometimes called the hardware or wetware — to distinguish these causal happenings from higher-scale structures, computations and so on[84,85].

In addition to the emphasis on lower-scale factors, the causal relationships in these mechanisms are described physically and mechanically, in terms of forces, action and motion. This is seen with rich causal verb descriptions of parts that bind, splice, activate, block, bend and so on, in contrast to simple causal language which states that one factor "causes" another outcome[86]. These neural mechanisms are sometimes analogized to machines in ways that highlight their physically interacting causal parts. As can be seen in these features and the analogy to machines, this notion of mechanism bears similarity to the original use developed by seventeenth-century mechanical philosophers.

A final main feature of this narrow notion of mechanism is that it requires scientists to provide substantial causal detail about the system. Gaining mechanistic understanding requires large amounts of detail, which leads scientists to claim that mechanisms are only ever partly identified as there is always more to uncover. This expansive character is distinct from that of abstract causal representations, such as network, dynamical and topological models, which are sparse and omit large amounts of information. In this usage, network models can capture causal relationships, but they do not reflect narrow mechanisms, which involve causal details that underpin, instantiate and capture inner workings of the system. This mechanism notion is not just reductive in character — it is often paired with reductive assumptions about neuroscientific explanation. These reductive assumptions include views that all phenomena in neuroscience are best explained with information from lower scales and that including increasing amounts of detail improves explanatory power.

The second use of mechanism is broad as it refers to a wide class of causal systems with different features. In some cases, this use of mechanism is so broad that it is simply synonymous with 'causation' and applies to nearly any type of causal system. In this usage, any system with causal relationships counts as a mechanism, even if these relationships do not include lower-scale factors or physical-impact, mechanical connections. Examples of this usage include "network mechanisms," "large-scale mechanisms", "systems-level mechanisms", "circuit mechanisms", "global mechanisms", "top-down mechanisms" and all lower-scale mechanisms referenced in the first usage[3,8–13,20,87–89]. This meaning of mechanism is non-reductive in the sense that it does not require that causal relationships or causal systems contain information about lower-scale parts and physical-impact interactions. For example, this usage is found in reference to large-scale causal relationships that span distant brain areas as, for example, instantiated by the notion of effective connectivity, lesion experiments, optogenetics, brain mapping techniques, various dynamic causal models and many others, as they reveal causal connections without lower-scale or physical detail[90,91] (Fig. 1). Consider the distinction between knowing that brain area A causes outcome B and knowing the causal details of how

this outcome is produced. This broad mechanism notion encompasses both the detailed, mechanical how and the higher-scale, abstract causal relationships.

As an example of this broad notion of mechanism, consider claims that in neuroscience, whatever outcome one is interested in, "the essential goal is to understand the mechanisms involved … [but] … we must be extremely careful to distinguish different specific kinds of mechanistic goals"[92]. In some cases, these goals include providing a detailed description of the "actual physical and chemical processes that underlie" the outcome in question. However, in other cases, these goals include "looking for something more general … [in which] it is not the details themselves that matter but the principles that are embodied in these details"[92]. This second claim suggests that mechanistic information is consistent with searching for more general, abstract principles "without incorporating all of the underlying biophysical details"[92]. In this manner, mechanistic enquiry is associated with causal information that can be detailed or abstract, lower scale or higher scale, and less constrained than the first conception suggests.

Further illustrations of this second usage are flexible mechanism definitions and interchanging mechanism language with causal language. As a first example, consider how "mechanistic" research is defined by the National Institutes of Health (NIH), which commonly uses this language in their grant calls. In these calls, the "NIH defines a mechanistic clinical trial as a study 'designed to understand a biological or behavioral process, the pathophysiology of a disease, or the mechanism of action of an intervention'"[93]. This definition is clearly broad and underspecified, able to encompass many types of causal information. Another example is present in neuroscientific studies that analyse "causal inferences", "causation" and "causal interactions" but, in the same discussion, switch usage to "mechanistic inferences", "mechanism" and "mechanistic understanding" without explanation[85,94]. The assumption is that 'causation' and 'mechanism' are synonymous — 'mechanism' is simply a filler word for 'causal system' and has little meaning beyond this. These papers sometimes define mechanism broadly as a "causal chain of events", suggesting that systems with multiple (two or more) causes qualify as mechanisms, no matter what their other features[85]. In other work it is suggested that the dominant notion of causality in neuroscience is mechanistic, in which the brain is understood as a "mechanistic system"[95,96].

A third use of 'mechanism' applies to models, systems and concepts that are not causal. Two examples of a non-causal use of 'mechanism' include using this term in reference, first, to various topological mechanisms and, second, to entities that are correlates, constitutive or realizers of a macro-scale system. With respect to the former, consider that models can contain helpful descriptions, warrant predictions and explain mathematically, all without containing causal information. As an example, consider reference to "topological mechanisms" — in some of these cases the explanations are powered by mathematical dependencies, as opposed to causal dependency relations[97]. Although well-known mathematical, non-causal explanations exist (such as in the Königsberg bridge case), their explanatory power is supplied by mathematical facts, proofs and theorems, whereas causal explanations require empirical information about interventionist difference-making relations[98,99]. Second, with respect to the latter non-causal use, consider that 'mechanism' is sometimes used synonymously with the 'correlates' and 'realizers' of a higher-scale system. However, what is not always appreciated is that notions of correlates, realizers and instantiation refer to relationships of constitution and not causation. To say that higher-scale outcome Y is constituted by X does not imply that Y is caused by X. In both of these non-causal cases, the 'mechanism' concept is used to refer to important types of relationship, but there are good reasons to resist claims that these relationships are causal.

In the order presented, these three uses of 'mechanism' increase in their breadth and flexibility. The first use applies to particular causal systems, the second extends to nearly all causal systems and the third broadens to include models and structures that are non-causal. In evaluating reasons for the more expansive notions of mechanism, it is worth considering the high status of the term in the field. It may be that the status of the term encourages use when a model is causal, or explanatory or important — yet these sorts of model are distinct, and they should be untangled.

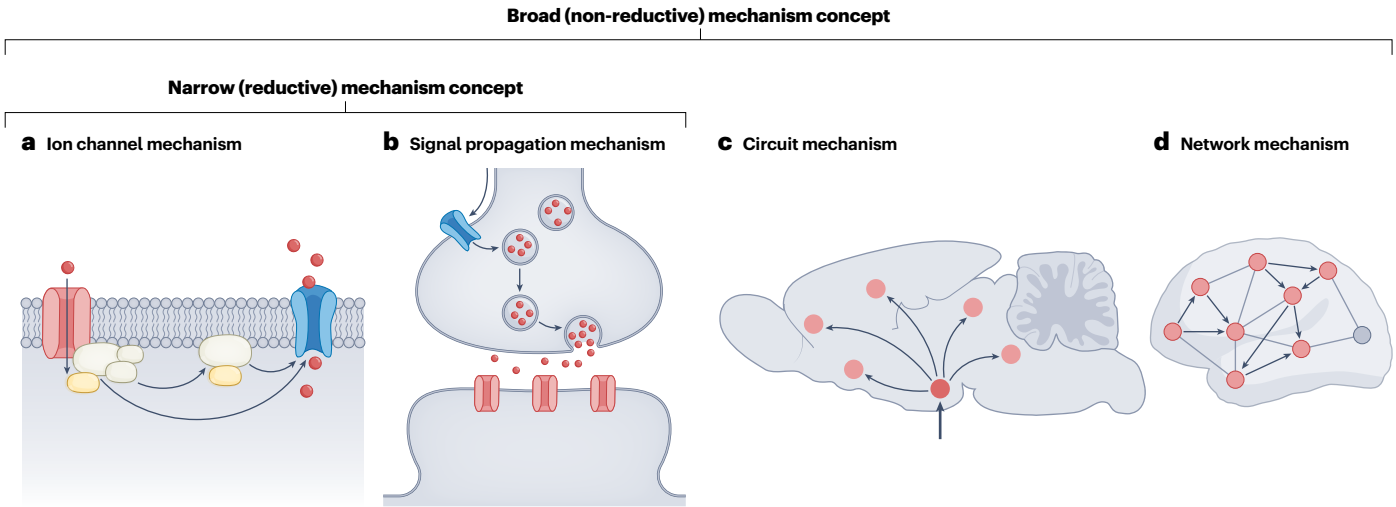

Fig. 1 | The narrow and broad notions of mechanism. a,**b**, The narrow notion of mechanism is found in examples such as ion channel mechanisms (panel **a**), cellular mechanisms and the mechanisms of signal propagation (panel **b**)[1,24,55,141]. **c**,**d**, The broader notion of mechanism is found in appeals to circuit mechanisms (panel **c**)[15,56,57,87], topological mechanisms, network mechanisms (panel **d**)[12,60,88] and cognitive mechanisms, but it includes systems with lower scale detail as well (panels **a** and **b**). The broad notion includes both abstract and fine-grain detailed causal systems as it is much more flexible about what counts as mechanistic.

Glossary

Causation
The relationship between cause and effect, which provides information about explanation and control (unlike relationships that are merely correlational).

Classification
Any organization or ordering of entities into categories.

Description
Any specification of a system or its features, which need not convey causal information.

Explanation
An answer to scientific 'why' questions, which often cites causes or causal structure.

Philosophy of science
An area of philosophy that concerns the methods, reasoning, concepts and foundations of science.

Prediction
An estimate of future outcomes or forecasting, which need not require causal information.

What are the challenges and implications of this variable usage?

In the field of neuroscience, the mechanism concept is associated with providing deep understanding and explanation of the brain. This association is seen in current publications, influential textbooks in various subfields of neuroscience and journals' descriptions of what they will (and will not) publish (Table 1). Further evidence is found in grant calls and funding solicitations for new research, which often specify their support for "mechanistic research" that uncovers the "mechanisms" of brain health and disease (Table 2). In these contexts, mechanisms are commonly tied to expressions of the main goals of the field and cited as the "fundamental", "foundational" and "basic" unit for understanding the brain[7,100–103].

It is obvious that the mechanism concept is viewed as central to neuroscience and in understanding the brain. However, the fact that this concept is defined in different ways throughout the field — and that there is often no clarification of which definition is intended — can engender considerable ambiguity about the fundamental goals, orientation and principles of the field. Whereas it is common for scientific fields to vigorously debate the exact meaning of important concepts (for example, species or laws of nature), a main difference here is that the variable usage of 'mechanism' remains largely unrecognized and unacknowledged in the field. This inattention to a fundamental term that is multiply defined leads to several challenges in the field; we outline and discuss five of these challenges below.

Lack of clarity

A first challenge of the variable meanings of mechanism is that this variability contributes to a substantial lack of clarity in the field, as mechanism is open to many interpretations. For example, when the field's overarching goals are detailed, funding inquiries are distributed, research is designed and implemented, and research findings are presented, the use of 'mechanism' can be understood in different ways and does not precisely communicate what is meant. In each case, use of this term requires further specification of its meaning — without such specification, the range of potential interpretations is great, from causal to non-causal systems. One example of this lack of clarity is found in the peer review process for neuroscience journals, in which there is considerable debate among reviewers and authors about whether the submitted paper has provided "mechanistic insights"[104–107], where this standard is often included in the journal's submission guidelines as a requirement for acceptance (Table 1). As is well documented in public reviews, these debates propose different understandings of what it takes to have mechanistic information — this can involve showing 'how' a causal process unfolds, providing evidence of manipulation or intervention, identifying factors that are necessary and sufficient for an outcome or providing a "mechanism of action" for the target of interest[104,108]. In neuroscience and other life science fields, it is often acknowledged that although providing mechanistic insight is a standard of the field, this standard "has also been cause for much uncertainty, mostly pertaining to what actually constitutes mechanistic insight"[109–111]. Although there is confusion about what mechanism means, the lack of mechanism information is a common reason for rejection of papers and grants, and many forms of advice for publication and grant submission explicitly suggest justifying why or how mechanism information is present[109]. The three mechanism categories we provide can help to orient discussions of what type of mechanism information is expected, required or missing from the work. Of course, this issue concerning lack of clarity is especially problematic owing to the fundamental status and common usage of 'mechanism' in the field. If mechanism is important to the field and our understanding of the brain, then a coherent, clear, and well-justified definition should be supplied.

Miscommunication

Another challenge is that the various meanings of mechanism encourage miscommunication. This can happen when one meaning is intended but another is assumed or when it is simply unclear which meaning is meant. One example of miscommunication is when mechanism language is used to refer to neural correlates of interest, such as the neural correlates of consciousness, and it is not clear whether mechanism is intended in a causal sense or in a non-causal, 'mere' correlation sense (which is suggested by the term 'correlate')[112–115]. This usage has engendered confusion about whether a mere correlate or cause of consciousness is being referred to, which is supported by slippage between a causal and a non-causal use of mechanism[116,117]. In other cases, funding calls for mechanistic research may be intended to support enquiry into higher-scale causal structures in the brain, but researchers may incorrectly interpret them as requiring the study of lower-scale, biophysical causes. Additionally, scientists may communicate their higher-scale mechanistic findings to experts and the public, with these audiences incorrectly assuming that biophysical mechanisms have been identified. Further misleading claims will result when mechanism is used in a non-causal sense but audiences incorrectly interpret the word mechanism in a causal manner. Making sure that experts do not talk past each other, that they accurately convey findings, ideas and strategies in moving forward, and that they properly inform the public requires more careful use than standard discussions have involved. Given that clear communication about causation is essential for the field — and that this is already challenging — the myriad meanings of mechanism further muddy the water[118].

Reductive assumptions

A third challenge of the many uses of mechanism is that the term can unintentionally support reductive assumptions. Consider that the first narrow, reductive meaning of mechanism is often assumed by default and that mechanism language often suggests an ease of manipulation, machine-like qualities and lower-scale detail that may

not be present[6,7,119–121]. When the broader notion is used to describe the foundations of the field, many audiences may interpret the word mechanism narrowly as relating to the claim that genuine causal understanding in neuroscience is reductive, always requiring molecular, cellular and other lower-scale factors. In much neuroscience literature there are suggestions that full, complete or deep explanations require "mechanistic" or causal insights from lower scales[119]. In other words, "it is often assumed that processes operating at the micro level are most mechanistically relevant"[119,122]. Alternatively, higher-scale causal structures are viewed as more superficial, as lacking full or complete causal understanding[123]. This view raises similar questions for science communication and the "responsible use of language", owing to the metaphors and meanings invoked[120]. The mechanism concept may misleadingly suggest that a causal model contains lower-scale detail or that such details give superior understanding in all contexts. It also makes it harder to acknowledge, account for and include higher-scale social and environmental causes into explanations of brain outcomes[124,125]. A central debate in neuroscience concerns how reductive our explanations should be — are lower-scale details always more explanatory or do higher-scale, abstract causal models sometimes provide better understanding? Switching between mechanism definitions can appear to support the narrow reductive view, even when this view is not intended or supported by the field.

Confusion about causation

A fourth challenge is that the many meanings of mechanism confuse discussions and understandings of causality. An example here is the use of Granger causality in elucidating the 'mechanisms' of brain outcomes[122,126]. Insofar as an interest in the mechanisms of brain health and disease is motivated by an interest in "changing" and "controlling" these outcomes, non-causal meanings of mechanism mislead because although they supply information about prediction, they do not supply information about control[94,127,128]. This distinction has led to many debates about how causality should be understood and whether it is well captured by Granger causality. The centrality of causation in neuroscience is undenied, as is the need for clarity regarding the notion(s) of causality that matter for the field, the methods that can identify it and proper ways to communicate about these topics[118]. The existence of many mechanism concepts threatens clarity here because they can conflate causation and mechanism and discourage discussions about causation by dropping causal language in favour of 'mechanism'. Mechanism is a concept that depends on causation — conflating these two concepts confuses attempts to clarify the causal standards of the field and how 'causation' should be understood, defined and communicated to various audiences. We are not suggesting that clear discussions about causality will result in immediate consensus on challenging topics (such topics have been debated in philosophy and science for centuries, at least). However, progress is hindered by the use of language that clouds the complex connection between mechanism and causation.

Undesirable incentives

A fifth challenge includes undesirable incentives and potential for misuse of the term. Like it or not, 'mechanism' has become a status term in the field — it often signals that a researcher is working on something important, something with causal information or something that deserves attention for funding, publication and citation. The high status of this concept can encourage scientists to use it liberally and with less care[120]. In particular, the considerable emphasis on mechanism can produce other undesirable incentives. On the one hand, if a scientist is operating with the narrow, reductive notion of mechanism, this makes it easy to dismiss theorizing that does not have this feature (such as dynamical systems, computational and other approaches that focus on 'upper-level' generalities) as not elucidating mechanisms and, hence, as unworthy. On the other hand, if a scientist is doing this type of non-reductive work, it can be very tempting to argue that it is mechanistic after all, in some extended sense of mechanism[94,123]. This can be an undesirable dialectic as the focus should be more directly on the strengths and weaknesses of the work itself, and not on whether it is correct to call it mechanistic. This dialectic is seen in reviewer–author discussions in peer review[104–107] and various papers in the field[119,123]. In many cases this dialectic can be improved by focusing on what type of causal information is sought and whether the scientific methods in use supply such information, and by having a clear sense of what type(s) of causal structure mechanism refers to.

This list of challenges is not exhaustive, and includes topics that are interrelated. In the next section, we explore how to address these challenges and provide suggestions on what should be done about the myriad definitions of mechanism.

How can we address these challenges?

Given the variable meanings of mechanism and the challenges that these variable meanings produce, how should the field proceed? A first major suggestion is that when mechanism is used, its meaning should be specified. Efforts to clarify usage will support communication in day-to-day research, grant inquiries and submissions, interdisciplinary exchanges and advancing theoretical discussions in the field. As long as mechanism is a fundamental term — perhaps the most commonly used causal concept in neuroscience — its variable meanings can encourage confusion unless the intended use is specified.

Is there more to say about how mechanism should be used in neuroscience? One way to make progress here is to consider the function of the concept — what is it intended to convey and communicate? Insofar as this term is causal, the third usage is clearly disadvantageous. This third notion does not support the aim of conveying causal information, although it can misleadingly suggest that it does. One worry is that the high status of the mechanism concept can encourage its use for models that are genuinely explanatory or important for other reasons, but nevertheless non-causal. This issue reflects an opportunity to clarify different types of explanation in the field and different models that provide understanding of the brain[64,94,98,129]. If mechanism is to have a clear causal meaning, then stretching it to include non-causal systems severely reduces the function and meaning of the term.

The second meaning of mechanism — applied to nearly any causal system — can serve the function of differentiating causal systems from non-causal ones. However, it cannot do much more. This usage cannot distinguish types of causal system, and makes the concept less meaningful by reducing it to any generic 'causal system'. This reduction is problematic insofar as we often view mechanism as communicating more than just causality — whether the content communicated is biophysical information, fine-grained detail or intermediates that specify 'how' an outcome is produced. If this concept just means 'causality' or 'causal system', why do neuroscientists continue to use the term or any of the rich, varied causal concepts that are common in the field? Examining different causal concepts in neuroscience — pathways, cascades, circuits, networks and so on — suggests that different terms are often used to refer to different types of causal systems. If this suggestion is correct, then collapsing all causal concepts and causal systems into

'mechanism' can make these distinctions harder to capture and glosses over varieties of causal explanation in the field.

The first notion of mechanism, which refers to causal systems with particular features, has many advantages. This meaning is often the default expected one; it has historical precedence, and it serves the function of communicating specific information about causal systems. This use makes sense of analogies to mechanical and machine-like systems, assumptions of reductive causal information and standard uses of the concept in other domains (including other sciences and everyday life). These advantages follow from a clear, constrained and unambiguous notion of mechanism. The ability to clearly capture diverse types of causal system and higher-scale causal structures is important – this is supported by a clear, distinct meaning of mechanism and exploring other important causal concepts in the field. Adopting this usage would not imply that all causal explanations in neuroscience should cite lower-scale details – instead, it would create room for other macro-scale, abstract and distinct causal systems that also explain. This approach supports a view in which "the brain can be interrogated, understood, and modeled at different levels … the explanatory or descriptive framework most appropriate for one level may not be the best approximation for another"[95]. In other words, gaining understanding and explanation of the brain does not privilege or prioritize any one particular scale of detail over another.

Now, why not adopt pluralism here and retain all three meanings of mechanism? Although pluralistic and umbrella concepts can be useful in neuroscience, there are reasons not to adopt the pluralist stance with respect to mechanism. Many pluralistic concepts pick out distinct classes of entity that all "cohere" in some way[130,131]. One challenge for this pluralistic path is that the distinct meanings of mechanism conflict with one another and lack basic coherence – each meaning contains assumptions that are incompatible with the others. If mechanism is a causal concept, then the third usage is inappropriate. If mechanism means something beyond mere 'causality' or 'causal system', then the second usage is overly flexible. Additionally, although pluralism has advantages in keeping frameworks open when "much remains to be learned about the variety and complexity" of systems in the world, this advantage is far less applicable to the mechanism case[130]. This is because mechanism is primarily used to convey and communicate pre-existing information about causality in the world and in research. Although it is important to remain open about different types of causality in the brain and methods that can be used to uncover them, clearly communicating in current work requires having terminology that supports clear, unambiguous statements. Although pluralism can be acceptable when supported by principled reasons, in other cases it may be a pseudo-solution when what is needed is clarity, careful deliberation and clear thinking about the theoretical underpinnings of the field.

Causation in neuroscience: mechanism and its alternatives

Although the term 'mechanism' is common in neuroscience, it should be placed in the context of the many other causal concepts used in the field[44,132]. Neuroscience uses a rich causal language, probably motivated by the study and discovery of diverse types of cause, causal relationship and causal structure in the brain and nervous system. Examples of this rich language include reference to the following: pathways, cascades and circuits; causes that control, trigger, constrain and predispose; and various causal topologies and networks with unique causal connections. Notions such as mechanism, pathway, cascade and circuit evoke different causal analogies that appear to convey information characteristic of distinct types of causal system. To say that a causal structure is most similar to a machine, or a roadway, or an amplifying cascade or an electrical circuit is to say different things about the system and to highlight the fact that causal systems differ with respect to their features. In fact, one way to appreciate causal complexity in the brain involves attending to distinct causal structures that are needed to explain different explanatory targets.

Aside from diverse causal systems, the mechanism concept should be examined in relation to questions about causality in neuroscience. Causal understanding is often associated with fundamental goals of the field – we should be sensitive to the possibility that mechanism may obscure and hinder these discussions. Key questions in this area are as follows:

- What definition (or definitions) of causality best capture usage in the field?
- How should causation be understood in neuroscience?
- Within a definition, what types of cause and causal structure are common?
- Does including lower-scale causal information always improve explanation and understanding?
- What standards do legitimate causal (and non-causal) explanations need to meet?
- Do answers to these questions distinguish causal models from models that are merely predictive, descriptive and classificatory?

These questions are theoretical, philosophical and practical. Answering them matters because they specify the types of causal information that scientists should study, they guide the types of work that grant agencies should fund and they aim to capture the field's view on what provides deep understanding and explanation of the brain. Answers to these questions matter for supporting communication about causality among experts and in how these experts communicate their work to the public. Finally, these questions matter for local inquiries about the features, payoffs and use of models, and for field-general inquiries about the principles, methods and strategies that support successful understanding in the field. By raising these questions, we do not assume that there are clear, single answers that will generate field-wide consensus. The point is to support clear discussion of these questions, and one helpful starting place is examining the various meanings of mechanism.

Published online: 11 January 2024

References

1. Woodward, J. in *Explanation and Integration in Mind and Brain Science* (ed. Kaplan, M.) 70–100 (Oxford Univ. Press, 2017).
2. Marinescu, I. E., Lawlor, P. N. & Kording, K. P. Quasi-experimental causality in neuroscience and behavioural research. *Nat. Hum. Behav.* **2**, 891–898 (2018).
3. Weichwald, S. & Peters, J. Causality in cognitive neuroscience: concepts, challenges, and distributional robustness. *J. Cog. Neurosci.* **33**, 226–247 (2021).
4. Uddin, L. Q. Cognitive and behavioural flexibility: neural mechanisms and clinical considerations. *Nat. Rev. Neurosci.* **22**, 167–179 (2021).
5. Crockett, M. J. & Fehr, E. Social brains on drugs: tools for neuromodulation in social neuroscience. *Soc. Cogn. Affect. Neurosci.* **9**, 250–254 (2014).
   This paper provides an important review of methodological considerations for the use of neuromodulation in social neuroscience.
6. Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, 2001).
7. Trappenberg, T. *Fundamentals of Computational Neuroscience* 2nd edn (Oxford Univ. Press, 2010).
8. Sporns, O., Tononi, G. & Kötter, R. The human connectome: a structural description of the human brain. *PLoS Computational Biol.* **1**, e42 (2005).

9. Cole, M. W., Ito, T., Bassett, D. S. & Schultz, D. H. Activity flow over resting-state networks shapes cognitive task activations. *Nat. Neurosci.* **19**, 1718–1726 (2016).
10. Pezawas, L. et al. 5-HTTLPR polymorphism impacts human cingulate–amygdala interactions: a genetic susceptibility mechanism for depression. *Nat. Neurosci.* **8**, 828–834 (2005).
11. Masse, N. Y., Yang, G. R., Song, H. F., Wang, X. & Freedman, D. J. Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nat. Neurosci.* **22**, 1159–1167 (2019).
12. Rubin, J., Rinzel, J., Arbib, M. A. & Bonaiuto, J. J. in *From Neuron to Cognition via Computational Neuroscience* (eds Arbib, M. A. & Bonaiuto, J. J.) 101–128 (MIT Press, 2016).
13. Vierling-Claasen, N. & Jones, S. Neural rhythms. in *From Neuron to Cognition via Computational Neuroscience* (eds Arbib, M.A. & Bonaiuto, J.J.) 129–158 (MIT Press, 2016).
14. Cabral, J., Kringelbach, M. L. & Deco, G. Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: models and mechanisms. *NeuroImage* **160**, 84–96 (2017).
15. Leaver, A. M., Espinoza, R., Wade, B. & Narr, K. L. Parsing the network mechanisms of electroconvulsive therapy. *Biol. Psychiatry* **92**, 193–203 (2022).
16. Salgado-Puga, K. & Pena-Ortega, F. Cellular and network mechanisms underlying memory impairment induced by amyloid β protein. *Protein Pept. Lett.* **22**, 303–321 (2015).
17. Cognigni, P., Felsenberg, J. & Waddell, S. Do the right thing: neural network mechanisms of memory formation, expression and update in *Drosophila*. *Curr. Opin. Neurobiol.* **49**, 51–58 (2018).
18. Parkin, B. L., Hellyer, P. J., Leech, R. & Hampshire, A. Dynamic network mechanisms of relational integration. *J. Neurosci.* **35**, 7660–7673 (2015).
19. Kedia, G., Mussweiler, T. & Linden, D. E. J. Brain mechanisms of social comparison and their influence on the reward system. *NeuroReport* **25**, 1255–1265 (2014).
20. Sanchez-Romero, R. & Cole, M. W. Combining multiple functional connectivity methods to improve causal inferences. *J. Cogn. Neurosci.* **33**, 180–194 (2021).
21. Schoonover, C. E., Fink, A. J. P., Poo, C. & Yuan, Q. Editorial: neuromodulation in olfaction, volume II. *Front. Cell. Neurosci.* **17**, 1146770 (2023).
22. Popovitchenko, T. & Rasin, M. Transcriptional and post-transcriptional mechanisms of the development of neocortical lamination. *Front. Neuroanat.* **11**, 102 (2017).
23. Saunders, N. R., Ek, C. J., Habgood, M. D. & Dziegielewska, K. M. Barriers in the brain: a renaissance? *Trends Neurosci.* **31**, 279–286 (2008).
24. Bechtel, W. &. Richardson, R. C. *Discovering Complexity* (MIT Press, 2010).
25. Craver, C. F. *Explaining the Brain* (Oxford Univ. Press, 2007).
26. Ross, L. N. Cascade verus mechanism: the diversity of causal structure in science. *Brit. J. Philos. Sci.* https://doi.org/10.1086/723623 (2023).
27. Woodward, J. *Making Things Happen* (Oxford Univ. Press, 2003).
This book presents a foundational and mainstream account of causation and causal explanation.
28. McIntosh, A. R. & Jirsa, V. K. The hidden repertoire of brain dynamics and dysfunction. *Netw. Neurosci.* **3**, 994–100 (2019).
29. Potochnik. P. *Idealization and the Aims of Science* (Univ. of Chicago Press, 2017).
30. Barwich, A. Imaging the living brain: an argument for ruthless reductionism from olfactory neurobiology. *J. Theor. Biol.* **512**, 110560 (2021).
31. Bickle, J., Mandik, P. & Landreth, A. The philosophy of neuroscience. *Standford Encylopedia of Philosophy* https://plato.stanford.edu/entries/neuroscience/ (2019).
32. Delehanty, M. The changing landscape of the philosophy of medicine. *Philos. Compass* https://doi.org/10.1111/phc3.12612 (2019).
33. Reiss, J. & Ankeny, R. Philosophy of medicine. *Standford Encylopedia of Philosophy* https://plato.stanford.edu/entries/medicine/ (2016).
34. Ward, Z. B. Cognitive variation: the philosophical landscape. *Philos. Compass* https://doi.org/10.1111/phc3.12882 (2022).
35. Henry, J. *The Scientific Revolution and the Origins of Modern Science* 2nd edn [Studies in European History] (Palgrave, 2001).
36. Dijksterhuis, E. J. *The Mechaniation of the World Picture* (Princeton Univ. Press, 1986).
37. Weber, M. *Philosophy of Developmental Biology* 1st edn (Cambridge Univ. Press, 2022).
38. Machamer, P., Darden, L. & Craver, C. F. Thinking about mechanisms. *Philos. Sci.* **67**, 1–25 (2000).
This paper contains an influential account of mechanism, which has inspired and influenced much current work on the topic.
39. Craver, C. F. & Darden, L. *In Search of Mechanisms* (Univ. Chicago Press, 2013).
40. Woodward, J. & Ross, L. Scientific explanation. *Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entries/scientific-explanation/ (2021).
41. Ross, L. & Woodward, J. Causal approaches to explanation. *Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entries/causal-explanation-science/ (2023).
42. Woodward, J. What is a mechanism? A counterfactual account. *Philos. Sci.* **69**, S366–S377 (2002).
43. Craver, C. F. & Tabery, J. Mechanisms in science. *Standford Encylopedia of Philosophy* https://plato.stanford.edu/entries/science-mechanisms/ (2015).
44. Ross, L. N. Causal concepts in biology: how pathways differ from mechanisms and why it matters. *Br. J. Philos. Sci.* **72**, 131–158 (2021).
This paper provides a clear account of the pathway and mechanism concepts, including how they differ and how they explain.
45. Russo, F. & Williamson, J. Interpreting causality in the health sciences. *Int. Stud. Philos. Sci.* **21**, 157–170 (2007).
46. Woodward, J. Mechanistic explanation: its scope and limits. *Proc. Aristot. Soc.* **LXXXVII**, 39–65 (2013).
47. Dupré, J. A. Living causes. *Aristot. Soc. Suppl. Vol.* **87**, 19–37 (2013).
48. Ross, L. N. Tracers in neuroscience: causation, constraints, and connectivity. *Synthese* **199**, 4077–4095 (2021).
49. Ross, L. N. Causal selection and the pathway concept. *Philos. Sci.* **85**, 551–572 (2018).
50. Gessell, B. S., Stanley, M. L., Geib, B. & De Brigard, F. in *Neural Mechanisms: New Challenges in the Philosophy of Neuroscience* (eds. Calzavarini, F. & Viola, M.) 35–55 (Springer, 2020).
51. Skillings, D. J. Mechanistic explanation of biological processes. *Philos. Sci.* **82**, 1139–1151 (2015).
52. Halina, M. in *The Routledge Handbook of Mechanisms and Mechanical Philosophy* (eds. Glennan, S & Illari, P.) 213–224 (Routledge, 2018).
53. MacLeod, M. & Nersessian, N. J. Modeling complexity: cognitive constraints and computational model-building in integrative systems biology. *Hist. Philos. Life Sci.* **40**, 17 (2018).
54. Dupré, J. & Nicholson, D. J. (eds.) *Everything Flows: Towards a Processual Philosophy of Biology* (Oxford Univ. Press, 2018).
This book expands on causal process views of living systems, which are contrasted with the mechanism perspective.
55. Dupré, J. A. Causally powerful processes. *Synthese* **199**, 10667–10683 (2021).
56. Bechtel, W. & Levy, A. Abstraction and the organization of mechanisms. Technical report. *Philos. Sci.* **80**, 241–261 (2013).
57. Robins, S. K. & Craver, C. J. in *The Oxford Handbook of Philosophy and Neuroscience* (ed. Bickle, J.) 41–67 (Oxford Univ. Press, 2009).
58. Kaplan, D. M. *Explanation and Integration in Mind and Brain Science* (Oxford Univ. Press, 2017).
59. Hedström, P. & Ylikoski, P. Causal mechanisms in the social sciences. *Annu. Rev. Sociol.* **36**, 49–67 (2010).
60. Ross, L. N. What is social structural explanation? A causal account. *Nous* https://doi.org/10.1111/nous.12446 (2023).
61. Hendry, D. F. The encompassing implications of feedback versus feedforward mechanisms in econometrics. *Oxf. Economic Pap.* **40**, 132–149 (1988).
62. Qin, Z. et al. Combustion chemistry of propane: a case study of detailed reaction mechanism optimization. *Proc. Combust. Inst.* **28**, 1663–1669 (2000).
63. Godfrey-Smith, P. in *The Oxford Handbook of Causation* (eds Beebee, H., Menzies, P. & Hitchcock, C.) 326–337 (Oxford Univ. Press, 2010).
64. Chirimuuta, M. Explanation in computational neuroscience: causal and non-causal. *Br. J. Philos. Sci.* **69**, 849–880 (2018).
65. Chirimuuta, M. Minimal models and canonical neural computations: the distinctness of computational explanation in neuroscience. *Synthese* **191**, 127–153 (2014).
66. Rust, N. C. & LeDoux, J. E. The tricky business of defining brain functions. *Trends Neurosci.* **46**, 3–4 (2023).
This paper outlines important challenges associated with defining, measuring and studying scientific phenomena of interest.
67. Rabuffo, G., Fousek, J., Bernard, C. & Jirsa, V. Neuronal cascades shape whole-brain functional dynamics at rest. *eNeuro* https://doi.org/10.1523/ENEURO.0283-21.2021 (2021).
68. Karuza, E. A., Thompson-Schill, S. L. & Bassett, D. S. Local patterns to global architectures: influences of network topology on human learning. *Trends Cogn. Sci.* **20**, 629–640 (2016).
69. Siddiqi, S. H., Kording, K. P., Parvizi, J. & Fox, M. D. Causal mapping of human brain function. *Nat. Rev. Neurosci.* **23**, 361–375 (2022).
70. Sharpley, C. F. & Bitsika, V. Joining the dots: neurobiological links in a functional analysis of depression. *Behav. Brain Funct.* **6**, 73 (2010).
71. Viding, E. & McCrory, E. Disruptive behavior disorders: the challenge of delineating mechanisms in the face of heterogeneity. *Am. J. Psychiatry* **177**, 811–817 (2020).
This article provides an important account of types of causal heterogeneity and causal complexity in psychiatry, psychology and neuroscience.
72. Thomas, M. S. C. & Coecke, S. Associations between socioeconomic status, cognition, and brain structure: evaluating potential causal pathways through mechanistic models of development. *Cogn. Sci.* **47**, e13217 (2023).
This paper presents an important current framework for understanding social causes, with reliance on the causal pathway concept.
73. Elsabbagh, M. Linking risk factors and outcomes in autism spectrum disorder: is there evidence for resilience? *BMJ* **368**, l6880 (2020).
74. Bielczyk, N. Z. et al. Disentangling causal webs in the brain using functional magnetic resonance imaging: a review of current approaches. *Netw. Neurosci.* **3**, 237–273 (2019).
75. Deng, Y., Reinhart, R. M., Choi, I. & Shinn-Cunningham, B. G. Causal links between parietal α activity and spatial auditory attention. *eLife* **8**, e51184 (2019).
76. Perrin, E. & Venance, L. Bridging the gap between striatal plasticity and learning. *Curr. Opin. Neurobiol.* **54**, 104–112 (2019).
77. Kuntsi, J. & Klein, C. in *Behavioral Neuroscience of Attention Deficit Hyperactivity Disorder and Its Treatment* Vol. 9 (eds Stanford, C. & Tannock, R.) 67–91 (Springer, 2012).
78. Willmore, L., Cameron, C., Yang, J., Witten, I. B. & Falkner, A. L. Behavioural and dopaminergic signatures of resilience. *Nature* **611**, 124–132 (2022).
79. Vemuri, P. et al. Amyloid, vascular, and resilience pathways associated with cognitive aging. *Ann. Neurol.* **86**, 866–877 (2019).
80. Sommer, I. E., Kleijer, H. & Hugdahl, K. Toward personalized treatment of hallucinations. *Curr. Opin. Psychiatry* **31**, 237–245 (2018).
81. Al-Diwani, A. A. J., Pollak, J. A., Irani, S. R. & Lennox, B. R. Psychosis: an autoimmune disease? *Immunology* **152**, 388–401 (2017).

82. Lappin, J. M. & Sara, G. E. Psychostimulant use and the brain. *Addiction* **114**, 2065–2077 (2019).
83. Estes, M. L. & McAllister, A. K. Maternal immune activation: implications for neuropsychiatric disorders. *Science* **353**, 772–777 (2016).
84. Srivastava, P. et al. Models of communication and control for brain networks: distinctions, convergence, and future outlook. *Netw. Neurosci.* **4**, 1122–1159 (2020).
85. Reid, A. T. et al. Advancing functional connectivity research from association to causation. *Nat. Neurosci.* **22**, 1751–1760 (2019).
86. Cartwright, N. Causation: one word, many things. *Philos. Sci.* **71**, 805–819 (2004).
87. Diano, S. A new brain circuit in feeding control. *Science* **361**, 29–30 (2018).
88. Langdon, C., Genkin, M. & Engel, T. A. A unifying perspective on neural manifolds and circuits for cognition. *Nat. Rev. Neurosci.* **24**, 363–377 (2023).
89. Havlík, M., Hlinka, J., Klírová, M., Adámek, P. & Horáček, J. Towards causal mechanisms of consciousness through focused transcranial brain stimulation. *Neurosci. Conscious.* **2023**, niad008 (2023).
90. Friston, K. J., Harrison, L. & Penny, W. Dynamic causal modelling. *NeuroImage* **19**, 1273–1302 (2003).
91. Misić, B., Mills, T., Taylor, M. J. & McIntosh, A. R. Brain noise is task dependent and region specific. *J. Neurophysiol.* **104**, 2667–2676 (2010).
92. O'Reilly, R. & Munakata, Y. *Computational Explorations in Cognitive Neuroscience* (MIT Press, 2000).
93. National Institute of Health. Support for clinical trials at NIMH. *National Institute of Mental Health* https://www.nimh.nih.gov/funding/opportunities-announcements/clinical-trials-foas (2023).
94. Mehler, D. & Kording, K. The lure of causal statements: rampant mis-inference of causality in estimated connectivity. Preprint at *arXiv* https://doi.org/10.48550/arXiv.1812.03363 (2018).
95. Wolff, S. B. & Olveczky, B. P. The promise and perils of causal circuit manipulations. *Curr. Opin. Neurobiol.* **49**, 84–94 (2018).
96. Faskowitz, J., Betzel, R. F. & Sporns, O. Edges in brain networks: contributions to models of structure and function. *Netw. Neurosci.* **6**, 1–28 (2021).
97. Tozzi, A. & Peters, J. F. A topological approach unveils system invariances and broken symmetries in the brain: system invariances and broken symmetries in the brain. *J. Neurosci. Res.* **94**, 351–365 (2016).
98. Woodward, J. in *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanation* (eds. Reutlinger, A. & Saatsi, J.) 117–137 (Oxford Univ. Press, 2019).
99. Ross, L. N. Distinguishing topological and causal explanation. *Synthese* **198**, 9803–9820 (2020).
100. Chiao, J. Y. Developmental aspects in cultural neuroscience. *Dev. Rev.* **50**, 77–89 (2018).
101. Amodio, D. M. The neuroscience of prejudice and stereotyping. *Nat. Rev. Neurosci.* **15**, 670–682 (2014).
102. Menon, V. & Uddin, L. Q. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* **214**, 655–667 (2010).
103. Falcon, M. I. et al. Functional mechanisms of recovery after chronic stroke: modeling with the virtual brain. *eNeuro* https://doi.org/10.1523/ENEURO.0158-15.2016 (2016).
104. Adams, J. N., Maass, A., Harrison, T. M., Baker, S. L. & Jagust, W. Cortical tau deposition follows patterns of entorhinal functional connectivity in aging. *eLife* **8**, e49132 (2019).
105. Lempart, J. et al. Mechanistic insights into the protective roles of polyphosphate against amyloid cytotoxicity. *Life Sci. Alliance* **2**, e201900486 (2019).
106. Capone, C. et al. Mechanistic insights into a TIMP3-sensitive pathway constitutively engaged in the regulation of cerebral hemodynamics. *eLife* **5**, e17536 (2016).
107. Shuvaev, S. A., Tran, N. B., Stephenson-Jones, M., Li, B. & Koulakoc, A. Neural networks with motivation. *OpenReview* https://openreview.net/forum?id=BJUVCEYDB (2020).
108. Seals, D. R. Publishing particulars: part 3. General writing tips, editing, and responding to peer review. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **324**, R409–R424 (2023).
109. Andersen, O. S. Editorial practices, scientific impact, and scientific quality. *J. Gen. Physiol.* **131**, 1 (2008).
110. Pugh, E. N. Kudos to reviewers for the *JGP*: you make our science better. *J. Gen. Physiol.* **133**, 129–130 (2009).
111. Pugh, E. N. The legacy of Olaf Sparre Andersen and future directions of the *Journal of General Physiology*. *J. Gen. Physiol.* **132**, 1–3 (2008).
112. Crick, F. & Koch, C. Towards a neurobiological theory of consciousness. *Semin. Neurosci.* **2**, 263–275 (1990).
113. Chalmers, D. J. in *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (ed. Metzinger, T.) 17–39 (MIT Press, 2000).
114. Wu, W. The neuroscience of consciousness. *Standford Encylopedia of Philosophy* https://plato.stanford.edu/entries/consciousness-neuroscience/ (2018).
115. Velmans, M. *Understanding Consciousness* 2nd edn (Routledge, 2009).
116. De Graaf, T. A., Hsieh, P. & Sack, A. T. The 'correlates' in neural correlates of consciousness. *Neurosci. Biobehav. Rev.* **36**, 191–197 (2012).
117. Aru, J., Bachmann, T., Singer, W. & Melloni, L. Distilling the neural correlates of consciousness. *Neurosci. Biobehav. Rev.* **36**, 737–746 (2012).
118. Rust, N. C. et al. A call for more clarity around causality in neuroscience. *Trends Neurosci.* **45**, 654–655 (2022).
119. Mill, R. D., Ito, T. & Cole, M. W. From connectome to cognition: the search for mechanism in human functional brain networks. *NeuroImage* **160**, 124–139 (2017).
120. Kueffer, C. & Larson, B. M. H. Responsible use of language in scientific writing and science communication. *BioScience* **64**, 719–724 (2014).
121. Boudry, M. & Pigliucci, M. The mismeasure of machine: synthetic biology and the trouble with engineering metaphors. *Stud. Hist. Philos. Sci. Part. C: Stud. Hist. Philos. Biol. Biomed. Sci.* **44**, 660–668 (2013).
122. Li, F. et al. Investigating dynamic causal network with unified Granger causality analysis. *J. Neurosci. Methods* **383**, 109720 (2023).
123. Li, G. & Yap, P. From descriptive connectome to mechanistic connectome: generative modeling in functional magnetic resonance imaging analysis. *Front. Hum. Neurosci.* **16**, 940842 (2022).
124. Paraouty, N. et al. Sensory cortex plasticity supports auditory social learning. *Nat. Commun.* **14**, 5828 (2023).
125. Hackman, D. A., Farah, M. J. & Meaney, M. J. Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nat. Rev. Neurosci.* **11**, 651–659 (2010).
126. Soleimani, B. et al. NLGC: network localized Granger causality with application to MEG directional functional connectivity analysis. *NeuroImage* **260**, 119496 (2022).
127. Woodward, J. A functional account of causation; or, a defense of the legitimacy of causal thinking by reference to the only standard that matters: usefulness (as opposed to metaphysics or agreement with intuitive judgment). *Philos. Sci.* https://doi.org/10.1086/678313 (2014).
128. Danks, D. & Davis, I. Causal inference in cognitive neuroscience. *WIREs Cogn. Sci.* **14**, e1650 (2023).
129. Ross, L. N. Dynamical models and explanation in neuroscience. *Philos. Sci.* **82**, 32–54 (2015).
130. Allen, C. On (not) defining cognition. *Synthese* **194**, 4233–4249 (2017).
131. Ludwig, D. & Ruphy, S. Scientific pluralism. *Standford Encylopedia of Philosophy* https://plato.stanford.edu/entries/scientific-pluralism/ (2021).
132. Woodward, J. *Causation With a Human Face* (Oxford Univ. Press, 2021).
133. Author guidelines. *Ann. Neurol.* https://onlinelibrary.wiley.com/page/journal/15318249/homepage/forauthors.html (2023).
134. Guide for authors. *Biol. Psychiatry* https://www.biologicalpsychiatryjournal.com/article/S0006-3223(20)31560-2/fulltext (2020).
135. Author information pack. *Exp. Neurol.* https://www.sciencedirect.com/journal/experimental-neurology/publish/guide-for-authors (2023).
136. Information for authors. *Lancet Neurol.* https://www.thelancet.com/pb/assets/raw/Lancet/authors/tln-info-for-authors-1686637133557.pdf (2023).
137. National Institute of Health. Neuroscience research. *National Institute of Neurological Disorders and Stroke* https://www.ninds.nih.gov/current-research/research-funded-ninds/neuroscience-research (2023).
138. NSF. Cognitive neuroscience (CogNeuro). Important information for proposers. *National Science Foundation* https://new.nsf.gov/funding/opportunities/cognitive-neuroscience-cogneuro (2015).
139. NSF. Collaborative research in computational neuroscience (CRCNS). Important information for proposers. *National Science Foundation* https://new.nsf.gov/funding/opportunities/collaborative-research-computational-neuroscience (2020).
140. NSF. Neural systems. Division of Integrative Organismal Systems Core Programs. Important information for proposers. *National Science Foundation* https://new.nsf.gov/funding/opportunities/neural-systems-0 (2011).
141. Daniel, H., Levenes, C. & Crepel, F. Cellular mechanisms of cerebellar LTD. *Trends Neurosci.* **21**, 401–407 (1998).

Acknowledgements
The authors thank J. Woodward for helpful comments on this manuscript. This work was supported by an NSF Career Award and Templeton Foundation Grant to L.N.R.

Author contributions
The authors both contributed to all aspects of the article preparation.

Competing interests
The authors declare no competing interests.

Additional information
Peer review information *Nature Reviews Neuroscience* thanks K. Friston; S. Genon, who co-reviewed with C. Rathkopf; and M. Halina for their contribution to the peer review of this work.