nature reviews psychology

https://doi.org/10.1038/s44159-022-00120-5

Perspective

Check for updates

The fundamental importance of method to theory

Rick Dale[1]✉, Anne S. Warlaumont[1] & Kerri L. Johnson[1,2]

Abstract

Many domains of inquiry in psychology are concerned with rich and complex phenomena. At the same time, the field of psychology is grappling with how to improve research practices to address concerns with the scientific enterprise. In this Perspective, we argue that both of these challenges can be addressed by adopting a principle of methodological variety. According to this principle, developing a variety of methodological tools should be regarded as a scientific goal in itself, one that is critical for advancing scientific theory. To illustrate, we show how the study of language and communication requires varied methodologies, and that theory development proceeds, in part, by integrating disparate tools and designs. We argue that the importance of methodological variation and innovation runs deep, travelling alongside theory development to the core of the scientific enterprise. Finally, we highlight ongoing research agendas that might help to specify, quantify and model methodological variety and its implications.

Sections

[1]Department of Communication, University of California, Los Angeles, Los Angeles, CA, USA. [2]Department of Psychology, University of California, Los Angeles, Los Angeles, CA, USA. ✉e-mail: rdale@ucla.edu

Introduction

In the 1980s, an emerging community of philosophers argued that the philosophy of science focused too much on theory and not enough on scientific methodology[1]. Sometimes called the 'new experimentalism', their approach to the philosophy of science emphasizes the central role of experimental methods (Box 1) and, in general, the specific procedures and contexts of scientific investigation[2–4]. These new experimentalists considered method to be worthy of philosophical investigation itself, and argued that such investigation can inform our understanding of scientific knowledge. Analogously, in contemporary psychological science, extensive discussions of research methods have led to proposals for how several concerns with the scientific enterprise, such as publication bias and other questionable research practices, might be overcome[5].

Box 1

What is a 'method'?

There is not one single consensus definition of 'method'. Within new experimentalism, methods are considered very broadly, often with a focus on experimentation. A practical way of defining methods is to look to manuals describing how to report them. For example, the American Psychological Association's publication manual remains a prominent guide to structuring manuscripts for many disciplines. In its third edition, the perennial 'Methods' section was described as requiring "enough detail to permit an experienced investigator to replicate the study," and had subsections for subjects, apparatus and procedure. The entirety of the methods section summary spanned fewer than two pages[166]. By its sixth edition, this summary had expanded to almost four pages, and included subsections for participants, sampling procedures, sample size and power, measures and covariates, research design and experimental manipulations or interventions[167]. The general notion that methods require detail to enable replication remains, but it is clear that the concept of methods is a very broad one, encompassing operationalization of variables, instrumentation, design matters, scripts for participants and task preparation.

Thus, we argue that 'methods' refers to many concepts that vary in specificity. Methods might focus primarily on instrumentation as central to obtaining measurements and defining variables. But methods can also involve research format, such as features of factorial experimental design. The American Psychological Association includes participants and sampling procedures as methods. One could also argue that knowledge of the wider contextual factors is needed to ensure successful replication, such as precisely when and where data were obtained. A conceptual critique here is that the American Psychological Association guidelines are too broad and nonspecific. This critique is mitigated by one major heuristic: the matters described in the American Psychological Association guidelines for reporting methods are the procedural configuration required to replicate a given outcome. This also aligns closely to the concept of the nomological machine (see below).

In this paper, we consider 'methods' in this broadest sense in most places, but sometimes imply or specify something more specific, such as the roles of instrumentation or research design.

Scientific psychologists cannot be said to neglect method, for the obvious reason that they are practising scientists. However, a focus on methods and their implications might be especially important for psychology. Human behaviour is complicated. For example, our attention changes depending on who is around us, whether we think we are being watched and what we think others know[6–9]; decision-making can be influenced by perception, recent decisions and other situational factors[10–13]. When a psychological process is affected by many other factors — cognitive, social and situational — fully understanding this process might require integration of measurements and methods relevant to those factors. For example, discovering that the presence of another person influences attention suggests a combination of methods that measure attention (such as eye tracking) with methods that enable more naturalistic social tasks[9].

In this Perspective, we argue that to tackle this kind of complexity, developing a variety of methodological tools and techniques should be regarded as a scientific goal in itself, one that is critical for advancing scientific theory. Methodological variety could be encouraged as a virtue of scientific epistemology alongside other purported virtues, such as theoretical simplicity, predictive power and applicability[14,15]. First, we describe this principle of methodological variety in more general terms. Next, we illustrate how this principle might be critical for complex psychological phenomena, using the domains of language and communication as examples. Finally, we discuss the implications of this general principle by emphasizing its promises and challenges for evolving scientific practice and describe ongoing research that might further understanding of theory and method and how they are related.

Methodological variety

Many philosophers of science and theorists have highlighted epistemological diversity across the sciences[14,16–25], with a similar emphasis on the important role of methodology. Nancy Cartwright (who is often listed among the 'new experimentalists') proposed a concept she called the nomological machine[17] that illustrates this centrality of methodological practice. The phrase 'nomological machine' describes how scientists carefully arrange their world to render observations and measurements from it: "…a fixed (enough) arrangement of components, or factors, with stable (enough) capacities that in the right sort of stable (enough) environment will, with repeated operation, give rise to the kind of regular behaviour that we represent in our scientific laws"[17].

Cartwright's nomological machine metaphor is an apt one for psychology. Experimental psychologists exploit sharply tuned paradigms for yielding systematic observations. The psychologist's experimental treatments are like procedural engines that yield systematic observations. The term nomological means relating to laws or lawful descriptions. Cartwright saw these procedural engines as the source of theories: generalized descriptions of patterns discovered in those procedures and observations.

The importance of methodological variety in this context might seem intuitive. Yet, psychological science has a long and well reasoned tradition of focusing on particular measures and experimental designs. This focus can be critical to achieving some clarity and understanding under specific manipulations, and is an important basis for direct replication[26]. The focus on specific measures and paradigms is also easy to justify because it reins in the complexity of human behaviour. However, some have raised concerns that a particular experimental design or methodological approach can become the focus of a research programme and thereby displace the broader phenomenon it is meant to investigate[27,28]. In other words, psychology focuses on experimental effects, which become the key phenomena defining a field of study[28].

As an example, consider again the process of visual attention. The long-standing approach of presenting passive stimuli to participants has yielded highly replicable laboratory findings. The effects observed in the Stroop and flanker tasks are very robust and have been used for decades to study attention and related cognitive processes[29–33]. Despite a systematic body of work produced using these paradigms, however, they cannot capture some important properties of human attention. For example, the distribution of visual attention differs when stimuli are other social entities rather than visually displayed items. Experiments have shown that attention is shifted away from a social stimulus when that stimulus is another person present in a room, suggesting that the potential for social interaction may alter the distribution of attention[7]. This relationship between attention and social stimuli has great theoretical relevance because it suggests that accounts of cognitive processes should include ecologically relevant dimensions of stimuli[9]. These theoretical ideas about attention in particular are guided by novel methods that embed social stimuli in attentional tasks[34] or in designs that are more ecologically valid[35]. In Cartwright's terminology, these researchers have devised nomological machines that reveal new and surprising influences on human attention. Furthermore, this research suggests that classic theories of attention need to be extended or amended to account for these effects of social stimuli. Thus, new 'machinery' yields new directions for theoretical development, and new theoretical ideas might suggest new machinery.

The relationship between theory and method is rich, and one vehicle connecting them is methodological variety. The importance of variety can be further illustrated from studies on how humans perceive potential communication partners. For example, in one set of studies researchers first found that people are reluctant to categorize others as belonging to stigmatized groups (such as a sexual minority group)[36]. The researchers then hypothesized that these findings stem from interpersonally benevolent biases. If this theoretical proposal about social cognition is correct, it would mean that some interpersonal decisions are characterized by more deliberative processing (cognitive style), a relation to pro-social orientation (motivation) and the pursuit of disconfirming information (reasoning). Of course, each element in this proposition invites distinct methods. Indeed, to test the tenets of their hypothesis, the researchers leveraged computer-mouse tracking, drift rate diffusion modelling, reaction time tasks, manipulation and measurement of prosocial tendencies, and multiple tests of confirmation biases. In combination, the multi-method approach provided comprehensive evidence about the motivated cognitive processes that produced the initial pattern of results[37].

To help explain complex phenomena, such a variety of methodologies probably travels alongside theoretical progress. Theorists and philosophers of science have identified epistemological criteria for evaluating the promise of a scientific theory. There is no consensus on these matters, but common criteria include simplicity, generality and falsifiability[15]. Methodological variety could belong in this set as a viable epistemological criterion because variety offers support to other criteria. For example, a theory that has been fostered and fine-tuned using many distinct methods will probably have greater generality and increasing robustness to new tests. Methodological variety also increases confidence in a theory, because the robustness of a theory can be assessed in light of the diversity of predictions it can make[38]. This confidence is akin to the diversity effect in decision-making and reasoning: that people make stronger inferences when supported by diverse sources of evidence[39,40].

This principle of variety does not merely require that psychologists use many distinct methods. It frames this methodological variety as a scientific virtue itself, one worthy of pursuit alongside other virtues. When an investigation uncovers some unexplained aspect of a psychological phenomenon, the researchers might seek new methods to study it. When attention researchers discover (or theoretically predict) the importance of social contexts, they develop new designs or adapt methods and measures from other domains. This variety does not necessarily mean that theories become more complicated. Instead, novel methods might also facilitate a desired condition of theories: that theories ought to become only sufficiently complex to account for a gradually increasing set of conditions and circumstances that are recognized to be important for the phenomenon in question. Thus, psychological scientists should actively cultivate methodological variety as a critical means of theoretical progress.

Human language and communication

In this section, we illustrate the importance of variety in the study of particular psychological phenomena drawn from the domain of human language and communication. We selected this domain because language has figured centrally in the history of psychological science and presents substantial challenges of complexity (Box 2). Indeed, language is among the most complex human skills[41]. One way to describe the complexity of language is in terms of different levels of behaviour. When people use language, even in mere seconds of conversation, their brains actively process sounds, words, sentences, meanings and topics. When communicating with language, people must fuse these elements meaningfully and in a timely way[42–45]. Yet a scientific understanding of language is even more complicated than this. Researchers must also understand how these complex skills are learned during childhood (or adulthood), and the languages people use reflect a long cultural history and practices that have been sustained across many generations[46,47]. Language might be structured to facilitate learning and use by human minds[48]. These considerations raise questions about the complex psychological underpinnings of all these processes. Answering these questions and addressing this complexity necessitates a variety of research methods.

This need for a variety of research methods to study language is further highlighted by theoretical debate about what language is. Some scholars emphasize the speech signal and focus on understanding the anatomical and neural adaptations that support speech production in humans compared to other primates[49]. To others, language is best equated to the capacity for recursive syntax (forming more and more complex sentences from simple grammatical rules), because this is one of language's most alluring properties[50]. Still others argue that language is best defined as a collective phenomenon, understood only as something distributed over a population of cognitive agents[51]. These distinct qualities of whatever is being called 'language' all invoke distinct research techniques. To illustrate, we consider two important areas related to language and communication.

Development of speech and language

Language learning has been a central and broadly influential topic in psychology and its history[52–54]. Language learning itself is complex and multifaceted, even when considering only parts of the overall language-learning process[55–57]. Take the case of trying to understand how a child comes to produce their first meaningful words, a milestone that many

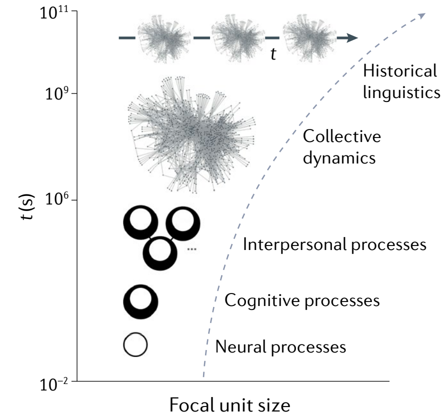
Box 2

Complexity and the law of requisite variety

A notion of methodological variety might also be derived from early history of the study of complex systems — in particular, cybernetics. Ashby[168,169] proposed that a self-regulating system's complexity or 'variety' will generally match or exceed the environmental challenges that system faces. This is often termed the law of requisite variety. This law has implications for cognitive organization. For example, a memory system might be structured so as to be in proportion to the complexity of the facts that need to be remembered[170], and a language's vocabulary will expand in response to the emergence of new things to talk about[171].

This concept can be applied to the design of science itself. It might be that the complexity of a phenomenon — its relevant processes and contexts at varying levels and goals of explanation — should induce a concomitant array of methodologies in order to tackle it. For example, language, a highly complex phenomenon, has its own explanatory substrata from neural processes in language to the collective phenomena of linguistic communities[172] (illustrated in the figure by approximate timescale *t*) exemplified not only in psychology but also in allied language sciences.

A critical next step is determining how to assess this complexity. This is relevant to assessing theories, such as the computational (or other) complexity that they imply. It is also relevant to assessing the complexity of tools on offer to test these theories. For example, assessing the computational complexity of cognitive theory has implications for the viability of these theories[137,173], including how they can be tested empirically[138,158]. If a theory can be shown to imply a computationally intractable process, then that theory would not be a viable explanatory candidate. This computational assessment has implications for understanding theory and how data bear upon it because the assessment checks the conditions required for a



cognitive theory to be feasible (namely, the computations should be tractable).

Similarly, the very wide variety of research goals across the sciences implies a rich kind of theoretical diversity. For example, some philosophers have proposed that 'explanatory styles' is a useful concept across the sciences. It is a phrase that describes the style of theorizing and associated research practices together, and it might invoke this kind of variety too[174]. An important next step is to codify these notions of theory and method, perhaps most promisingly by quantifying their characteristics and connections. Bibliometric tools now make this possible[139,175–177]. Importantly, this understanding of research style, including diverse methods, must be developed in a manner that sustains coherent theoretical developments and avoids mere accumulation of effects[20,28,151,178].

infants reach at around their first birthday. This accomplishment is partly the result of input from adults, including acoustic information that changes at timescales ranging from milliseconds to minutes or longer, as children learn the sounds that correspond to words. It is also partly the result of the infant's learned knowledge of their physical and social environments, such as their ability to identify and categorize the entities to which the word refers. And it is partly the result of a long process of vocal motor learning wherein infants learn through trial and error to control the muscles of their vocal tract with the larynx, lips, tongue, velum and jaw[56]. These biomechanical structures are also moving targets from a motor learning perspective: owing to dramatic physical growth and body shape change in an infant's first year, the precise actions an infant must execute to generate a particular sound type change as the infant grows.

Scientists need diverse methodologies to fully understand the phenomena of vocal learning and first-word production[58–62] (Fig. 1). One research strategy is to devise computational models that solve similar problems, but are simplified in such a way that psychologists can explore the relationships among multiple variables[63–65]. For example, neural network models are in very abstract terms statistical learning systems with properties that resemble a human brain — simulated neurons that connect and process information and learn from their input[66]. One computational study used a neural network linking motor control of a fluid dynamics model of the speech system with simulated bursts of dopamine that regulate learning rate. The dopamine bursts are intended to correspond with the receipt of social and intrinsic rewards that real infants experience when they babble[67,68] (Fig. 1a,b). This model's 'babble' becomes increasingly speech-like over time. This modelling approach extends our understanding of motor and learning subsystems relevant to first-word production, linking levels of analysis that are difficult to test in vivo.

This methodological approach is complemented by other methodological approaches involving human and other animal subjects. Numerous methodological advances over the past four decades have contributed to scientific progress on this front. A common approach from the mid-twentieth century onward involved time-restricted observational studies: a field researcher might go into the home or bring caregivers and parents into the laboratory[69], and record (and later transcribe) an hour or so of audio (and sometimes video) data capturing child–adult interactions. These data formed the basis for the groundbreaking online repository called Child Language Data Exchange System (CHILDES)[70]. CHILDES is a large-scale database of systematically coded and openly shared conversational data, foreshadowing by a decade or more the 'big data' approach to scientific research of the present day. CHILDES led to fundamental new insights about the role of caregiver input in language development. For example, analysis of the relationship between parental and child language use in the CHILDES data provided overwhelming evidence that language

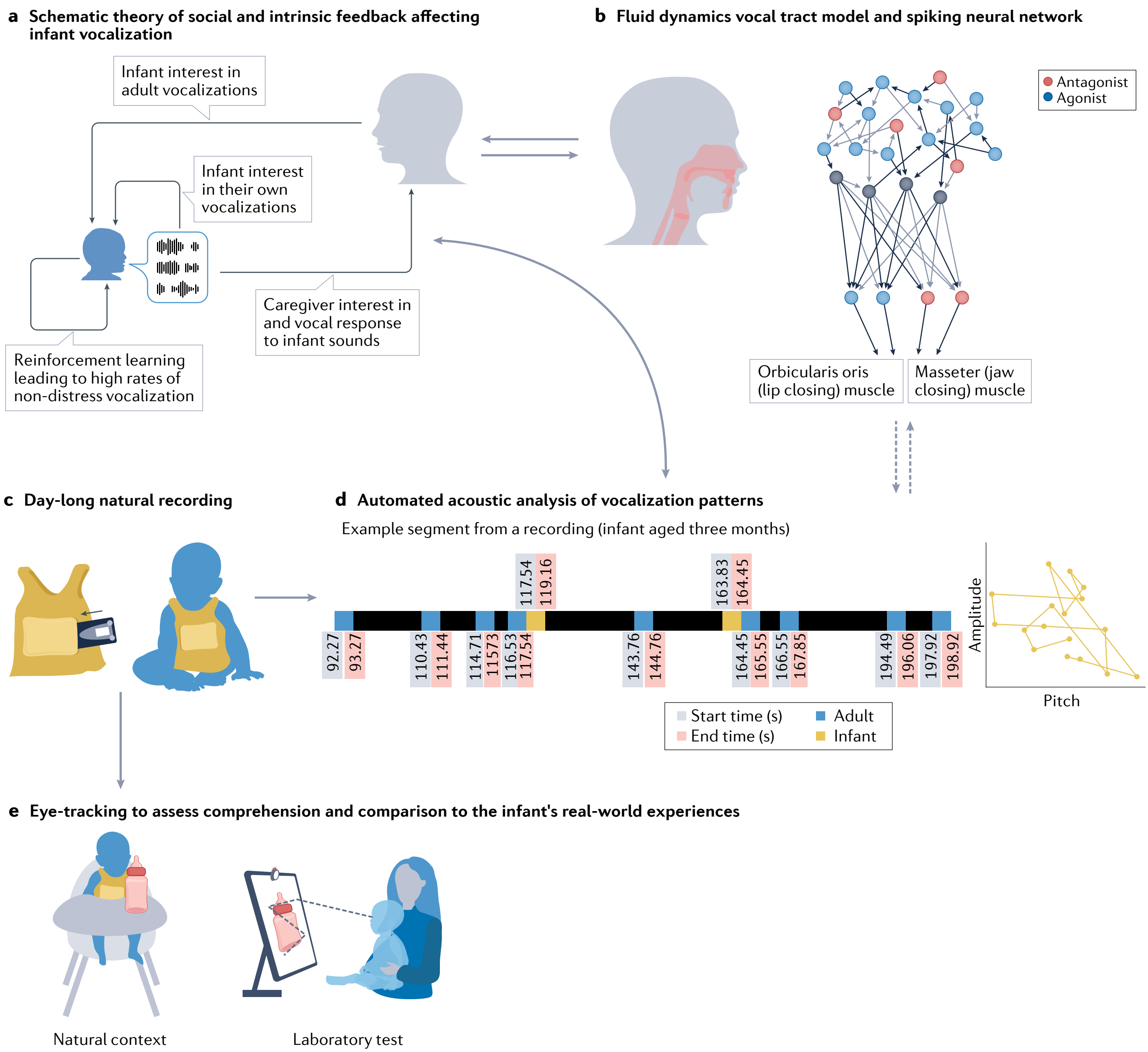


Fig. 1 | Interrelationships between methods used to study the emergence of speech. a, A schematic theory posits that infants and caregivers form a feedback circuit that promotes infants' production of more mature speech sounds. **b**, A computational model that includes both a spiking neural network model of motor activity and learning and a computational model of the muscle and fluid dynamics in the vocal tract permits more detailed exploration of potential mechanisms that could underlie the infant side of the feedback circuit in part **a**. Thus, schematic theories inform and are informed by computational models (solid arrows between parts **a** and **b**). **c**, Day-long audio recording is collected using recorders worn by the child. **d**, Analyses of data from part **c** support the testing of hypotheses about how social interaction influences infant vocal production (solid arrow between parts **a** and **d**) and should enable direct comparison with behaviour of computational models (dashed arrows between parts **b** and **d**). **e**, Day-long audio recording is also used to relate specific words that infants hear to the knowledge of sound–meaning mappings that they demonstrate using the looking-while-listening paradigm in a highly controlled experimental setting (solid arrow between parts **c** and **e**). This information in turn contributes to schematic theories of how infants learn words based on their linguistic and other sensory experiences. Part **b** is adapted from ref.[164], CC BY 4.0 and ref.[68], CC BY 4.0. Part **d** is adapted from ref.[83], CC BY 4.0.

learning is influenced by the frequency of the input to children, mainly that words and grammatical structures tend to be acquired earlier when they are more frequent in adult speech[71]. Computational studies have utilized the transcript data from CHILDES to demonstrate the feasibility of learning semantic and grammatical properties of words based on language samples that children have actually heard[72–75]. Such models have also informed and been informed by in-lab statistical learning experiments. In these studies, participants learn sequential stimuli designed to have certain properties of real languages[76–80]. These methods further test predictions from CHILDES analysis and computational models, and have supported an important role for the mind's basic learning processes.

Starting in the mid-2000s, researchers have gone beyond the classic approaches described above to measure infants' experiences at home. For example, small electronic recorders inserted into a child's clothing (Fig. 1c) can continuously capture the ambient auditory environment for up to 16 hours. Recoding longitudinally and across many children yields massive datasets of many thousands of hours of audio data, enabling researchers to study aspects of real-world auditory, vocal and social experiences that were previously unavailable. For instance, automated identification of child and adult vocalizations within these recordings (Fig. 1d) reveal that infant and adult vocalizations are clustered at multiple timescales throughout the day and that input types are distributed non-uniformly[81–84]. These features align infant communication with a suite of behaviours that can be usefully described as a kind of 'foraging'. This theoretical frame connotes a 'search' that infants are engaging in, exploring their vocal-articulatory landscape for regions of high saliency and social attention[83]. Such a theoretical framing highlights the similarities between the process of learning to speak and other search processes, such as retrieving items from memory, searching for food in the wilderness, and searching the internet for information. This suggests the possibility that the neural mechanisms and environmental considerations influencing those other types of search processes might also influence human infant speech and language development. This frame also suggests that patterns regarding the nesting of vocal behaviour across timescales could be more informative than measures that do not take into account nested timescales, as has been found in other domains.

Such day-long audio recordings can also be combined with advances in head-mounted cameras and infant eye-tracking experiments (as well as with long-used parental reports of infant vocabulary knowledge). This marks a new era of measuring infants' experiences at home in a manner that can be related to laboratory eye-tracking research on word understanding[85] and parent reports of the words infants produce (Fig. 1e). Research using this multi-method and multi-site approach has found that infants' knowledge of word sound–meaning mappings is related to the statistical structure of what they hear and see in the home[73,86,87].

In sum, increasingly diverse methodological approaches enable researchers to make different inroads into the very complex and multi-faceted set of processes involved in language learning and to advance theory by integrating these approaches.

Language in use

Human conversation relies on the real-time coordination of numerous verbal and nonverbal behaviours between interlocutors. Understanding how such spontaneous coordination occurs requires methods that range from highly structured laboratory experiments that establish elementary social and linguistic scenarios to more naturalistic observations that probe the relations between behavioural signals[88–91]. In both cases, novel methodologies have corroborated existing theory in new and sometimes unexpected ways and have enabled important theoretical developments.

For example, in the mid-1990s, eye-tracking methods led to groundbreaking discoveries about how the mind exploits extraneous factors when processing language[92]. Eye movements reveal how an individual interprets verbal instructions as they unfold. Upon hearing "look at the candle," listeners might briefly fix their attention on an image of a candy (a sweet) when it is present in their visual environment while the sound "can…" is being pronounced[93] (Fig. 2a, left). Thus, eye movements reveal that certain inferences are made in parallel while the sound signal in speech is being processed. In other words, the mind dynamically integrates multiple sources of information (sounds and sights), and does not wait for the word to be completed before making guesses about what the word will be, based on visual cues in the environment. This observation had major theoretical implications for language, primarily by supporting theories that frame language as a parallel[94,95] rather than serial[96] process. Eye-tracking methods have continued to open up important new avenues to arbitrate theories. For example, by tracing the eyes as they scan the visual world, researchers can test whether participants actively predict a word or object before it has even been mentioned[97,98]. It is even possible to reconstruct sentences spoken by someone by using statistical patterns in their eye movements[99].

A variant of this eye-tracking method uses the computer mouse[100]. In these tasks, participants see or hear a word or instruction, and move their mouse to an answer (Fig. 2a, right) while the x and y coordinates of their mouse cursor are recorded. Borrowing the example from above, when hearing 'click the candle', a slight perturbation in a respondent's arm movements might occur when a candy is also on the screen[101]. Similar perturbations in computer mouse trajectories are detectable for sounds, words and syntax[102]. Thus, like eye movements, subtle arm movements reveal that the mind integrates various pieces of information in parallel. Because these methods sample the dynamics of different external movement systems, they strengthen inferences about the cognitive processes for language that underlie both measures.

These types of method have revealed that perceived social variables – race, gender and social status – can profoundly affect interpersonal processing such as stereotyping and evaluative judgements even when such variables are not ostensibly related to the decision participants are making[103]. Precisely how social cognition impinges on language processing remains an exciting future direction for interdisciplinary scholarship. Combining innovative methods for exploring group perception[104,105] and organization[106,107] might be useful for testing how features of groups relate to signal production and comprehension[108,109].

Methods of dynamic measurement can also be combined with observation (Fig. 2b) to study language and communication in more naturalistic settings. For example, automated body motion analysis can be combined with speech segmentation analysis and computer vision[110–115]. Studies that take this approach to investigating language use and communication in context support the view that verbal behaviours (such as word choice) and nonverbal behaviours (such as body motion and eye movements) become interdependent when two or more people interact[43,44,116–118,119]. In related studies, analysing patterns of neural activity while participants experience longer and

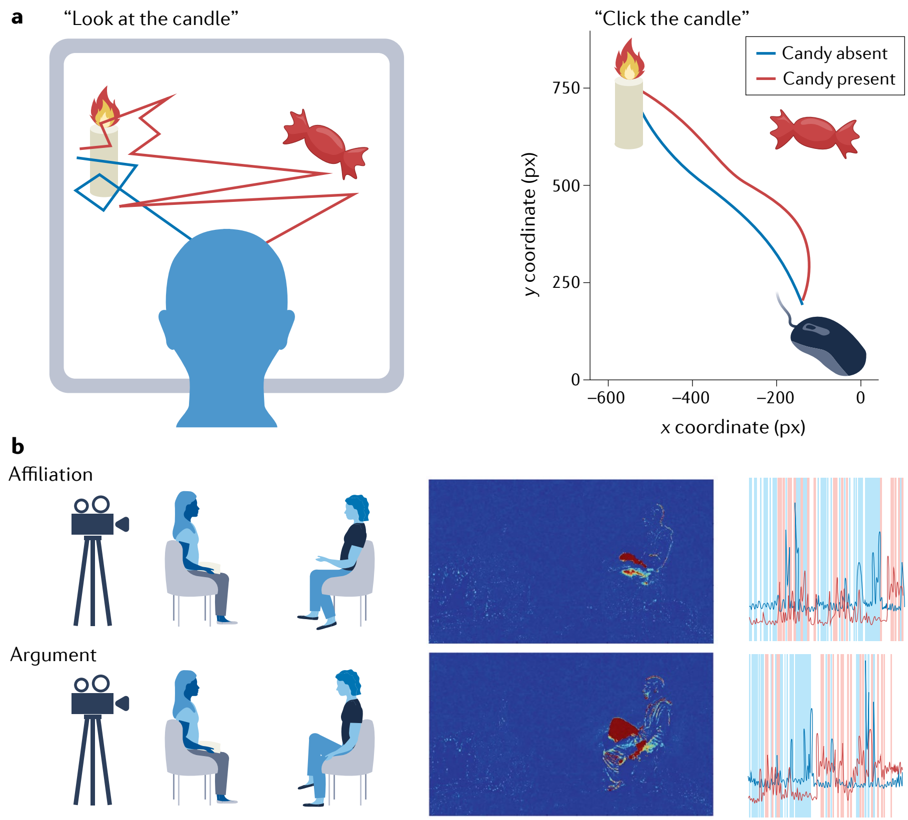


Fig. 2 | An illustration of methodological expansion to study language and communication. a, Eye movement studies show that participants temporarily look at a competing object ('candy') before a word ('candle') is fully spoken. Similar effects are observed using computer-mouse tracking, where arm movement trajectories reveal subtle competition when the word for an incorrect object ('candy') has feature overlap with the target object ('candle')[101,165]. **b**, Natural interactions can be established in the laboratory and dynamic measurement tools can capture time series of behaviour that underpin language and communication. For example, simple video analysis methods can be used to track body movement ('body language') and speech while participants interact. Using video- and audio-processing tools, time series can be extracted such as body (lines in the figures) and voice (bands in the figures) variables. Quantitative analyses can then be conducted to relate these dynamics to communication outcomes such as conflict[111].

uninterrupted samples of naturalistic stimuli may reveal how language is represented in the brain[120–122]. One study suggests that language might rely on a specialized brain network that actively integrates diverse forms of information, especially in the semantic domain[123]. Combining brain imagining and computational linguistics in this way also promises to expand our understanding of how language mediates between communication partners[121,124–126].

These new methods pose a new theoretical challenge. The rich multimodality of language, and its relentless processing during natural language use, demand a clearer understanding of how these modalities and other parts of language fit together in such a highly coordinated way[117]. Many of these methods generate new kinds of dataset, especially ones that challenge current analytic approaches. For example, very large-scale data and their dynamics are inviting new quantitative perspectives on language as a kind of multimodal coordinated performance – an ensemble cast of many cognitive processes, rather than a loose collection of discrete, distinct cognitive processes[127–129]. The ability to measure and analyse previously unmeasured large-scale and multidimensional data illustrates the potential for methodology to facilitate theoretical developments[130,131].

Theory and method as a coupled system

The previous sections used language and communication to illustrate that studying a complex psychological phenomenon requires a varied array of methodological tools. They also illustrate two kinds of methodological variety. In the case of language learning, wide-ranging methods can be integrated. Neural networks, behavioural experiments and body analysis serve as diverse means of targeting core theoretical questions. Such varied methods might be needed when certain manipulations are impossible – such as experimentally manipulating a child's long-term language experience. However, in language processing, methodological variety can come from translating existing methods more subtly. For example, tracking behavioural dynamics in different effectors (hands, eyes) helps to triangulate the underlying characteristics of cognitive processes.

These examples suggest that methodological development characterizes many research domains. However, questions remain about precisely how these methodologies should affect theory, and how theory should then fuel further methodological development. The evolution of theory and method together is likely to be subtle, with no strictly prescribed path (Box 3). Understanding methods and

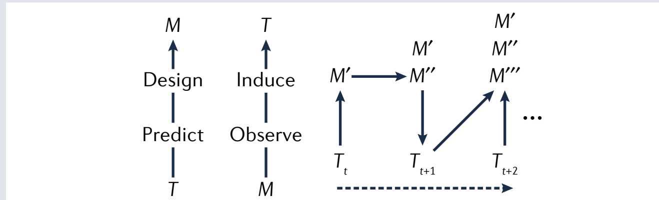
Box 3

Relations between method and theory

The new experimentalists were, in part, reacting to a famous edict in the philosophy of science that observation is theory-laden: that when taking measurements in a particular context, the manner in which these observations are interpreted and the impact they will have are filtered through a lens of prior theoretical commitments[179]. There is therefore no ‘raw observation’. Nevertheless, an argument that theory-independent experimentation is possible emerged, in the sense that raw measurement with an apparatus can be retained by scientists even if their theories change[1].

Theory and methods (for example, measurement and observation) are tightly interconnected in the sciences. In psychological science, the exhortation to develop clearer connections between experimental results and theory has had some formal development[141], and some have devised best practices and heuristics for building such theories[28,146,180]. Although we focus on methodological variety here, it is important to recognize this mutuality between theory and method.

But this suggests a chicken-and-egg problem. When a new scientific programme emerges, is it a new theory (*T*) that leads to new methods (*M*) and experimentation? Or is it a discovery in unrelated or theoretically unconstrained observation that leads to development of theory? It would seem that both of these paths (see figure, part **a**) are possible in practice. There might also be a subtler, temporally evolving relationship between theory and method (see figure, part **b**, where *t* represents time). A given theory can be modified over time, motivating new methods, and observations from these methods might lead to further methodological innovation. Growing



methodologies serve to support evolving theory, which motivates methodological innovation along the way.

However, this mutuality need not characterize every scientific advance. In a prominent account of the emergence of thermometry (the measurement of temperature), the philosopher and historian of science Hasok Chang describes a history of elaborate and iterative experimentation, without a robust anchoring theory. Such a theory would come only later to justify the various intricacies of this domain (such as boiling points)[181]. According to this historical account, the relationship between theory and measurement might be even more subtle, and can involve a more decoupled evolution. We therefore do not urge one stance on this particular question, nor do we suggest that there is a general normative position on whether and how method and theory should relate. The position we explore here would simply see methodological variety as a parameter in this dialogue, and one that has intriguing cognitive implications, especially for complex psychological phenomena.

methodological variety in the context of theory development might be a valuable topic of study in itself.

Some recent computational research, including by philosophers of science, offers examples of investigating scientific practice using computer models. These models typically render scientists and their practices as simplified entities that can be computationally specified and investigated[132]. For example, Zollman’s formal model of scientific induction illustrates how communities using varying practices lead to different epistemological outcomes[24]. The social connectivity among ‘researchers’ and how different methods are distributed are set up in a computer model to explore simulated scientific outcomes, such as convergence on one method among all researchers. An implication of this model is that the most efficient strategy might not be maximum variety at all times, but rather transient periods during which increased variety could facilitate progress. For example, groups should not converge too quickly on one methodology. In related research, Devezer and colleagues devise an agent-based model to explore how various aspects of the scientific process influence epistemological outcome[133]. One key finding from that work is that innovation accelerates convergence on the ‘truth’, and that epistemic diversity (characterized as ‘research strategy’) is a collective benefit to the scientific process.

An important related question is how methodological variety relates to broader questions about diversity in science. Diversity has been argued to be especially critical to some aspects of scientific process and discourse, such as specifying an initial research problem[134]. Here, too, computational modelling might contribute. For example, O’Connor and Bruner offer a model that investigates diversity among simulated scientific agents[135]. Using a game-theoretic framework, they show that majority–minority gradients, when present, can adversely influence small groups, and can deter members of a minority group from communicating with the majority. Importantly, these effects can emerge without the presence of a clear bias among majority agents. A general implication of this work is that sustaining diversity is not simply a matter of dampening a bias towards uniformity (whether in identity or in research strategy), because there might be emergent properties of coordination that lead to these effects on minority groups. (See ref.[136] for an overview of social factors involved in scientific practice from a philosophical perspective.)

Other quantitative techniques might further contribute to understanding the evolving terrain of theory and method. One approach is to explore the computational limitations of theory[137,138], and then to assess empirical implications for how research techniques might (or might not) address them. Bibliometric techniques might also offer insight into how methods are distributed across different theories. For example, in a bibliometric analysis of thousands of cognitive science abstracts, the theoretical inclination of a publication could be predicted by its network of terms, many of which were methodological[139].

This computational and quantitative research suggests that method and theory have systematic linkages. The social distribution of scientific practice, and the collective benefit of various community strategies, might have a sharp effect on the overall epistemological outcome. Put simply, theory and method might be a kind of coupled system.

Challenges and promises

In addition to reinforcing and coupling to existing theory, novel methods can introduce new ideas, new concepts, new tethers to reality — new nomological machines. As such, theory might expand in new and unconventional ways in response to these new techniques. In this section we consider challenges and promises in considering the principle of methodological variety as more than 'just methods'[140].

A first challenge, and related to the prior section, is integrating theory and method. For example, some have argued that a theory crisis[141] is occurring alongside psychology's replication crisis[5,28,141–145], such that the development of theory is not given sufficient attention[28,146]. Without precisely tuned theoretical goals, it is unclear which studies should be replicated, and questions persist about the evidential weight replications should have[141]. Consequently, methodological novelty might find friction with existing validated measurements. New and unfamiliar methods might not immediately have the strongest links to general theories or specific models. These links have to be fostered through iterated study and validation, alongside the development and advancement of theory[146].

Another challenge is one of breadth. Science is a human enterprise, and it has human constraints[136]. One of these constraints is expertise. Methodological variety is tempered by the time and effort required for individual scientists and their laboratories to achieve methodological expertise. In the past couple of decades, psychologists have adopted the team science approach of other disciplines, such as biology and medicine, to overcome this issue[147,148]. Published research increasingly includes many authors with distinct contributions and expertise. This approach facilitates an epistemological division of labour in teams or in a laboratory. Indeed, understanding how scientist teams work together is a budding area of psychological research itself[149]. However, this team science approach comes with risks, as authorship ethics and credit assignment have wide implications, particularly for the career trajectories of early-career researchers[150].

There might be a virtuous mutuality between methodological variety and discussions about reproducibility and theory. This mutuality might help to overcome the challenge of breadth. For example, adopting transparent research practices enables the expertise of one scientist to be more easily transferred to others. A laboratory can more readily expand its methodological repertoire by making use of openly shared materials, code and data. Nevertheless, the idea that methodological expertise and theoretical understanding could be distributed might be difficult to reconcile with traditional notions of a scientist as an expert on one or a small set of specific paradigms.

Despite these challenges, encouraging methodological variety (and innovation) as one important goal might be a critical ingredient of theoretical advances, especially when novel methods facilitate exploration. For example, the focus on statistical hypothesis testing might sometimes be based on the flawed assumption that effects link rigorously to theoretical models. Instead, researchers could engage in more nonconfirmatory investigation of the complex relationships among measures and paradigms to facilitate clearer conditions for rigorous, reliable and replicable links between effects and theories (see ref.[151] in particular for discussion).

In addition, methodological variety encourages researchers to bridge theories and methods, especially ones that cut across laboratories and the theoretical debates in which they are engaged. In the domain of conversation and discourse, some have remarked that theoretical competition might be illusory, in part driven by disparate methods employed by distinct laboratories[88,128]. The principle of methodological variety puts in plain sight the limiting effects of focusing too closely on particular paradigms.

Similarly, there are strong incentives to fine-tune theories through the many sources of open data beyond one's own laboratory[62,152–154]. Open data can be considered a means of tapping into 'open methods', because datasets from other laboratories are often collected from distinct experimental designs, contexts and cultures. For example, large-scale datasets collected by linguists on the grammatical properties of language[155] have been fused with large-scale datasets of linguistic populations and cultures[156], enabling psychologists to explain how a language's grammar might relate to cognition and social structure[108,109].

Along similar lines of reasoning, a mixture of methods can help to triangulate a given theoretical construct, and help to distil its relationship to other variables of interest. At least since MacCorquodale and Meehl's[157] classic contrast of intervening variables and hypothetical constructs, many domains of psychology have struggled to isolate constructs statistically and by experimentation[158]. For example, perhaps 'language' is simply the coordination of many 'non-linguistic' things (such as perception and attention, memory, and sequential learning)[159]. Conversely, language could be best characterized by processes of symbolic grammar or syntax, reifying this term as a core psychological construct[160]. Diverse methods are critical to mitigating such debate about constructs by devising effective means of measuring them, or by assembling diverse designs and data with which to triangulate them[151].

Conclusion

The scientific method has served as a well known metaphor for infant and child learning: infants and children test hypotheses by engaging their world[161]. Moreover, like scientists, babies have multiple approaches to learning. For example, when a baby explores a specific physical object, they might hold it and rotate it along multiple axes[162]; they might mouth the object[163], drop it, bang it, and/or throw it; they might attempt to give or show it to another person; they might also ask a caregiver what it is, or attempt to name it or describe it in their own words. As the baby engages with the object in these ways, they develop an understanding of the object that is not reducible to a simple verbal description or a single visual prototype. Rather, they acquire neurally encoded knowledge about the multimodal features of the object, its affordances for action, its linguistic label or labels and their grammatical properties, and social-pragmatic information about the object. This might be considered analogous to a group of scientists actively pursuing diverse methodological approaches to tackle their own very complex environment of investigation. Scientific progress on many psychological questions is, after all, still in its infancy.

We suggest three promising avenues that might facilitate such exploratory behaviour. First, innovation requires scientists to exploit methods from adjacent fields and put them to different uses. Just as new uses for pharmaceutical treatments fundamentally alter the approved uses for a particular medication, repurposing a method common within one field for use in an allied but adjacent field can facilitate new discoveries and theory development. Additionally, such endeavours enhance interdisciplinary reach and, by extension, the theoretical innovation, of the work. Second, methodological variety could be distributed across and subsequently shared between laboratories within a discipline. By establishing a diverse methodological toolkit within a field, the broader science stands to benefit by growing a collective knowledge base that will ultimately be codified in overarching theory. Finally, methodological variety could be enhanced within individual laboratories. Adopting varied methods to interrogate hypotheses that stem from one's theory could sharpen insights, providing a methodological dexterity that facilitates discovery. Thus, researchers might benefit from encouraging active and deliberate exaptation of methods from allied fields, embracing the diversified approaches of others within their own subdiscipline, and adopting a multiple-method approach within their own laboratory.

A sibling journal to this one, *Nature Methods*, launched over a decade ago with an editorial paean to methods, proclaiming that methods should be recognized as fundamental to scientific progress — that methods papers should be granted a similar appreciation to papers that simply apply them[140]. The editorial lamented that technical creativity in this regard is often not rewarded in its own right, and consequently there are limited incentives for developing and describing novel techniques. Here we have argued that the importance of methods runs deep, travelling alongside theory development to the core of the scientific enterprise. Methods are so critical to scientific progress that stating as much is intuitive. But a potential principle of methodological variety is distinct from this basic idea. Evidential diversity is often framed retrospectively: that we have a theory in hand, and we then assess it for evidential diversity[38]. A principle of methodological variety would urge researchers to think in proactive terms, to seek out distinct scenarios or novel measures that might further advance their theories by refining them, connecting them to other domains, and putting them to new and challenging tests. Researchers seek simplicity and specificity for their theories, and reproducibility and replication of their predicted effects. Theories could be further enriched and strengthened through methodological variety.

Published online: 29 November 2022

References

1. Hacking, I. *Representing and Intervening: Introductory Topics in The Philosophy of Natural Science* (Cambridge Univ. Press, 1983).
2. Mayo, D. G. in *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* Vol. 1994, 270–279 (Cambridge Univ. Press, 1994).
3. Ackermann, R. The new experimentalism. *Brit. J. Philos. Sci.* **40**, 185–190 (1989).
4. Simons, M. & Vagelli, M. Were experiments ever neglected? Ian Hacking and the history of philosophy of experiment. *Phil. Inq.* **9**, 167–188 (2021).
5. Nosek, B. A. et al. Promoting an open research culture. *Science* **348**, 1422–1425 (2015).
6. Richardson, D. C., Dale, R. & Tomlinson, J. M. Conversation, gaze coordination, and beliefs about visual context. *Cogn. Sci.* **33**, 1468–1482 (2009).
7. Laidlaw, K. E., Foulsham, T., Kuhn, G. & Kingstone, A. Potential social interactions are important to social attention. *Proc. Natl Acad. Sci. USA* **108**, 5548–5553 (2011).
8. Richardson, D. C. et al. Joint perception: gaze and social context. *Front. Hum. Neurosci.* **6**, 194 (2012).
9. Risko, E. F., Richardson, D. C. & Kingstone, A. Breaking the fourth wall of cognitive science: real-world social attention and the dual function of gaze. *Curr. Dir. Psychol. Sci.* **25**, 70–74 (2016).
10. Levin, I. P., Schneider, S. L. & Gaeth, G. J. All frames are not created equal: a typology and critical analysis of framing effects. *Organ. Behav. Hum. Decis. Process.* **76**, 149–188 (1998).
11. Pothos, E. M. & Busemeyer, J. R. A quantum probability explanation for violations of 'rational' decision theory. *Proc. R. Soc. B* **276**, 2171–2178 (2009).
12. Pärnamets, P. et al. Biasing moral decisions by exploiting the dynamics of eye gaze. *Proc. Natl Acad. Sci. USA* **112**, 4170–4175 (2015).
13. Vinson, D. W., Dale, R. & Jones, M. N. Decision contamination in the wild: sequential dependencies in online review ratings. *Behav. Res. Methods* **51**, 1477–1484 (2019).
14. Longino, H. E. Gender, politics, and the theoretical virtues. *Synthese* **104**, 383–397 (1995).
15. Hansson, S. O. in *Stanford Encyclopedia of Philosophy* (ed. Zalta, E. N.) https://plato.stanford.edu/archives/fall2021/entries/pseudo-science/ (Stanford Univ., 2015).
16. Dupré, J. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science* (Harvard Univ. Press, 1993).
17. Cartwright, N. *The Dappled World: A Study of the Boundaries of Science* (Cambridge Univ. Press, 1999).
18. McCauley, R. N. & Bechtel, W. Explanatory pluralism and heuristic identity theory. *Theory Psychol.* **11**, 736–760 (2001).
19. Mitchell, S. D. Integrative pluralism. *Biol. Phil.* **17**, 55–70 (2002).
20. Abrahamsen, A. & Bechtel, W. in *Contemporary Debates in Cognitive Science* (ed. Stainton, R.) 159–187 (Blackwell, 2006).
21. Kellert, S. H., Longino, H. E. & Waters, C. K. *Scientific Pluralism* (Univ. Minnesota Press, 2006).
22. Dale, R., Dietrich, E. & Chemero, A. Explanatory pluralism in cognitive science. *Cogn. Sci.* **33**, 739–742 (2009).
23. Weisberg, M. & Muldoon, R. Epistemic landscapes and the division of cognitive labor. *Phil. Sci.* **76**, 225–252 (2009).
24. Zollman, K. J. S. The epistemic benefit of transient diversity. *Erkenntnis* **72**, 17–35 (2010).
25. Horst, S. Beyond reduction: from naturalism to cognitive pluralism. *Mind Matter* **12**, 197–244 (2014).
26. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* **349**, 943 (2015).
27. Lin, H., Werner, K. M. & Inzlicht, M. Promises and perils of experimentation: the mutual-internal-validity problem. *Perspect. Psychol. Sci.* **16**, 854–863 (2021).
28. van Rooij, I. & Baggio, G. Theory before the test: how to build high-verisimilitude explanatory theories in psychological science. *Perspect. Psychol. Sci.* **16**, 682–697 (2021).
29. MacLeod, C. M. The Stroop task: the gold standard of attentional measures. *J. Exp. Psychol. Gen.* **121**, 12–14 (1992).
30. Eriksen, C. W. The flankers task and response competition: a useful tool for investigating a variety of cognitive problems. *Vis. Cogn.* **2**, 101–118 (1995).
31. Dickter, C. L. & Bartholow, B. D. Ingroup categorization and response conflict: interactive effects of target race, flanker compatibility, and infrequency on N2 amplitude. *Psychophysiology* **47**, 596–601 (2010).
32. Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M. & Ferrand, L. The loci of Stroop effects: a critical review of methods and evidence for levels of processing contributing to color-word Stroop effects and the implications for the loci of attentional selection. *Psychol. Res.* **86**, 1029–1053 (2022).
33. Barzykowski, K., Wereszczyski, M., Hajdas, S. & Radel, R. Cognitive inhibition behavioral tasks in online and laboratory settings: data from Stroop, SART and Eriksen Flanker tasks. *Data Brief.* **43**, 108398 (2022).
34. Gobel, M. S., Kim, H. S. & Richardson, D. C. The dual function of social gaze. *Cognition* **136**, 359–364 (2015).
35. Gallup, A. C., Chong, A. & Couzin, I. D. The directional flow of visual information transfer between pedestrians. *Biol. Lett.* **8**, 520–522 (2012).
36. Lick, D. J. & Johnson, K. L. Straight until proven gay: a systematic bias toward straight categorizations in sexual orientation judgments. *J. Pers. Soc. Psychol.* **110**, 801 (2016).
37. Alt, N. P., Lick, D. J. & Johnson, K. L. The straight categorization bias: a motivated and altruistic reasoning account. *J. Pers. Soc. Psychol.* https://doi.org/10.1037/pspi0000232 (2020).
38. Popper, K. *Conjectures and Refutations: The Growth of Scientific Knowledge* (Routledge, 2002).
39. Heit, E. & Hahn, U. Diversity-based reasoning in children. *Cogn. Psychol.* **43**, 243–273 (2001).
40. Heit, E., Hahn, U. & Feeney, A. in *Categorization Inside and Outside the Laboratory: Essays in Honor of Douglas L. Medin* 87–99 (American Psychological Association, 2005).
41. MacWhinney, B. in *The Handbook of Linguistics* 2nd edn (eds. Aronoff, M. & Rees-Miller, J.) 397–413 (Wiley, 2017).
42. Stivers, T. et al. Universals and cultural variation in turn-taking in conversation. *Proc. Natl Acad. Sci. USA* **106**, 10587–10592 (2009).
43. Louwerse, M. M., Dale, R., Bard, E. G. & Jeuniaux, P. Behavior matching in multimodal communication is synchronized. *Cogn. Sci.* **36**, 1404–1426 (2012).
44. Fusaroli, R., Bjørndahl, J. S., Roepstorff, A. & Tylén, K. A heart for interaction: shared physiological dynamics and behavioral coordination in a collective, creative construction task. *J. Exp. Psychol. Hum. Percept. Perform.* **42**, 1297 (2016).
45. Rasenberg, M., Özyürek, A. & Dingemanse, M. Alignment in multimodal interaction: an integrative framework. *Cogn. Sci.* **44**, e12911 (2020).
46. Dunn, M., Greenhill, S. J., Levinson, S. C. & Gray, R. D. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* **473**, 79–82 (2011).
47. Hua, X., Greenhill, S. J., Cardillo, M., Schneemann, H. & Bromham, L. The ecological drivers of variation in global language diversity. *Nat. Commun.* **10**, 1–10 (2019).
48. Christiansen, M. H. & Chater, N. The now-or-never bottleneck: a fundamental constraint on language. *Behav. Brain Sci.* **39**, e62 (2016).
49. Fitch, W. T., De Boer, B., Mathur, N. & Ghazanfar, A. A. Monkey vocal tracts are speech-ready. *Sci. Adv.* **2**, e1600723 (2016).
50. Hauser, M. D., Chomsky, N. & Fitch, W. T. The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579 (2002).
51. Cowley, S. J. *Distributed Language* (John Benjamins, 2011).
52. Samuels, R. Nativism in cognitive science. *Mind Lang.* **17**, 233–265 (2002).
53. Behme, C. & Deacon, S. H. Language learning in infancy: does the empirical evidence support a domain specific language acquisition device? *Phil. Psychol.* **21**, 641–671 (2008).

54. Chomsky, N. *4. A Review Of BF Skinner's Verbal Behavior* (Harvard Univ. Press, 2013).
55. Vihman, M. M. *Phonological Development: The Origins of Language in the Child* (Blackwell, 1996).
56. Oller, D. K. *The Emergence of the Speech Capacity* (Psychology Press, 2000).
57. Clark, E. V. & Casillas, M. *First Language Acquisition* (Routledge, 2015).
58. Goldstein, M. H., King, A. P. & West, M. J. Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl Acad. Sci. USA* **100**, 8030–8035 (2003).
59. Warlaumont, A. S. Modeling the emergence of syllabic structure. *J. Phonet.* **53**, 61–65 (2015).
60. VanDam, M. et al. HomeBank: An online repository of daylong child-centered audio recordings. *Semin. Speech Lang.* **37**, 128–142 (2016).
61. Elmlinger, S. L., Schwade, J. A. & Goldstein, M. H. The ecology of prelinguistic vocal learning: parents simplify the structure of their speech in response to babbling. *J. Child Lang.* **46**, 998–1011 (2019).
62. Roy, B. C., Frank, M. C., DeCamp, P., Miller, M. & Roy, D. Predicting the birth of a spoken word. *Proc. Natl Acad. Sci. USA* **112**, 12663–12668 (2015).
63. McClelland, J. L. The place of modeling in cognitive science. *Top. Cogn. Sci.* **1**, 11–38 (2009).
64. Smaldino, P. E. in *Computational Social Psychology* (eds Vallacher', R., Read, S. J. & Nowak, A.) 311–331 (Routledge, 2017).
65. Guest, O. & Martin, A. E. How computational modeling can force theory building in psychological science. *Perspect. Psychol. Sci.* **16**, 789–802 (2021).
66. Elman, J. L., Bates, E. A. & Johnson, M. H. *Rethinking Innateness: A Connectionist Perspective on Development* Vol. 10 (MIT Press, 1996).
67. Warlaumont, A. S., Westermann, G., Buder, E. H. & Oller, D. K. Prespeech motor learning in a neural network using reinforcement. *Neural Netw.* **38**, 64–75 (2013).
68. Warlaumont, A. S. & Finnegan, M. K. Learning to produce syllabic speech sounds via reward-modulated neural plasticity. *PLoS One* **11**, e0145096 (2016).
69. MacWhinney, B. & Snow, C. The child language data exchange system: an update. *J. Child. Lang.* **17**, 457–472 (1990).
70. MacWhinney, B. *The CHILDES Project: Tools for Analyzing Talk* 3rd edn (Psychology Press, 2014).
71. Kachergis, G., Marchman, V. A. & Frank, M. C. Toward a "standard model" of early language learning. *Curr. Dir. Psychol. Sci.* **31**, 20–27 (2022).
72. Lewis, J. D. & Elman, J. L. A connectionist investigation of linguistic arguments from the poverty of the stimulus: learning the unlearnable. In *Proc. Annual Meeting of the Cognitive Science Society* Vol. 23, 552–557 (eds. Moore, J. D. & Stenning, K.) (2001).
73. Regier, T. & Gahl, S. Learning the unlearnable: the role of missing evidence. *Cognition* **93**, 147–155 (2004).
74. Reali, F. & Christiansen, M. H. Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cogn. Sci.* **29**, 1007–1028 (2005).
75. Foraker, S., Regier, T., Khetarpal, N., Perfors, A. & Tenenbaum, J. Indirect evidence and the poverty of the stimulus: the case of anaphoric *one*. *Cognit. Sci.* **33**, 287–300 (2009).
76. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).
77. McMurray, B. & Hollich, G. Core computational principles of language acquisition: can statistical learning do the job? *Dev. Sci.* **12**, 365–368 (2009).
78. Frost, R., Armstrong, B. C., Siegelman, N. & Christiansen, M. H. Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cogn. Sci.* **19**, 117–125 (2015).
79. Isbilen, E. S., Frost, R. L. A., Monaghan, P. & Christiansen, M. H. Statistically based chunking of nonadjacent dependencies. *J. Exp. Psychol. Gen.* https://doi.org/10.1037/xge0001207 (2022).
80. Ruba, A. L., Pollak, S. D. & Saffran, J. R. Acquiring complex communicative systems: Statistical learning of language and emotion. *Top. Cogn. Sci.* https://doi.org/10.1111/tops.12612 (2022).
81. Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S. & Kello, C. T. Multiple coordination patterns in infant and adult vocalizations. *Infancy* **22**, 514–539 (2017).
82. Mendoza, J. K. & Fausey, C. M. Everyday music in infancy. *Dev. Sci.* **24**, e13122 (2019).
83. Ritwika, V. et al. Exploratory dynamics of vocal foraging during infant-caregiver communication. *Sci. Rep.* **10**, 10469 (2020).
84. Mendoza, J. K. & Fausey, C. M. Quantifying everyday ecologies: principles for manual annotation of many hours of infants' lives. *Front. Psychol.* **12**, 710636 (2021).
85. Fernald, A., Zangl, R., Portillo, A. L. & Marchman, V. A. in *Developmental Psycholinguistics: On-line Methods in Children's Language Processing* (eds. Sekerina, I. A., Fernández, E. M. & Clahsen, H.) Vol. 44, 97 (John Benjamins, 2008).
86. Weisleder, A. & Fernald, A. Talking to children matters: early language experience strengthens processing and builds vocabulary. *Psychol. Sci.* **24**, 2143–2152 (2013).
87. Bergelson, E. & Aslin, R. N. Nature and origins of the lexicon in 6-mo-olds. *Proc. Natl Acad. Sci. USA* **114**, 12916–12921 (2017).
88. Brennan, S. E., Galati, A. & Kuhlen, A. K. in *Psychology of Learning and Motivation* (ed. Ross, B. H.) Vol. 53, 301–344 (Elsevier, 2010).
89. Streeck, J., Goodwin, C. & LeBaron, C. *Embodied Interaction: Language and Body in the Material World* (Cambridge Univ. Press, 2011).
90. Goodwin, C. *Co-operative Action* (Cambridge Univ. Press, 2017).
91. Dale, R., Spivey, M. J. in *Eye-Tracking In Interaction. Studies On The Role Of Eye Gaze In Dialogue* (eds. Oben, B. and Brône, G.) 67–90 (John Benjamins, 2018).
92. Richardson, D. C. & Spivey, M. J. in *Encyclopedia of Biomaterials and Biomedical Engineering* 573–582 (CRC Press, 2004).
93. Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. & Sedivy, J. C. Integration of visual and linguistic information in spoken language comprehension. *Science* **268**, 1632–1634 (1995).
94. Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M. & Sedivy, J. C. Eye movements and spoken language comprehension: effects of visual context on syntactic ambiguity resolution. *Cogn. Psychol.* **45**, 447–481 (2002).
95. Richardson, D. C., Dale, R. & Spivey, M. J. Eye movements in language and cognition. *Methods Cogn. Linguist.* **18**, 323–344 (2007).
96. Ferreira, F. & Clifton, C. Jr The independence of syntactic processing. *J. Mem. Lang.* **25**, 348–368 (1986).
97. Kamide, Y., Altmann, G. T. & Haywood, S. L. The time-course of prediction in incremental sentence processing: evidence from anticipatory eye movements. *J. Mem. Lang.* **49**, 133–156 (2003).
98. Coco, M. I., Keller, F. & Malcolm, G. L. Anticipation in real-world scenes: the role of visual context and visual memory. *Cogn. Sci.* **40**, 1995–2024 (2016).
99. Coco, M. I. & Keller, F. Scan patterns predict sentence production in the cross-modal processing of visual scenes. *Cogn. Sci.* **36**, 1204–1223 (2012).
100. Kieslich, P. J., Henninger, F., Wulff, D. U., Haslbeck, J. M. & Schulte-Mecklenbeck, M. in *A Handbook of Process Tracing Methods* 111–130 (Routledge, 2019).
101. Spivey, M. J., Grosjean, M. & Knoblich, G. Continuous attraction toward phonological competitors. *Proc. Natl Acad. Sci. USA* **102**, 10393–10398 (2005).
102. Freeman, J., Dale, R. & Farmer, T. Hand in motion reveals mind in motion. *Front. Psychol.* **2**, 59 (2011).
103. Freeman, J. B. & Johnson, K. L. More than meets the eye: split-second social perception. *Trends Cogn. Sci.* **20**, 362–374 (2016).
104. Goodale, B. M., Alt, N. P., Lick, D. J. & Johnson, K. L. Groups at a glance: perceivers infer social belonging in a group based on perceptual summaries of sex ratio. *J. Exp. Psychol. Gen.* **147**, 1660–1676 (2018).
105. Sneller, B. & Roberts, G. Why some behaviors spread while others don't: a laboratory simulation of dialect contact. *Cognition* **170**, 298–311 (2018).
106. Atkinson, M., Mills, G. J. & Smith, K. Social group effects on the emergence of communicative conventions and language complexity. *J. Lang. Evol.* **4**, 1–18 (2019).
107. Raviv, L., Meyer, A. & Lev-Ari, S. The role of social network structure in the emergence of linguistic structure. *Cogn. Sci.* **44**, e12876 (2020).
108. Lupyan, G. & Dale, R. Language structure is partly determined by social structure. *PLoS One* **5**, e8559 (2010).
109. Lupyan, G. & Dale, R. Why are there different languages? The role of adaptation in linguistic diversity. *Trends Cogn. Sci.* **20**, 649–660 (2016).
110. Wu, L., Waber, B. N., Aral, S., Brynjolfsson, E. & Pentland, A. Mining face-to-face interaction networks using sociometric badges: predicting productivity in an IT configuration task. *Inf. Syst. Behav. Soc. Methods* https://doi.org/10.2139/ssrn.1130251 (2008).
111. Paxton, A. & Dale, R. Argument disrupts interpersonal synchrony. *Q. J. Exp. Psychol.* **66**, 2092–2102 (2013).
112. Alviar, C., Dale, R. & Galati, A. Complex communication dynamics: exploring the structure of an academic talk. *Cogn. Sci.* **43**, e12718 (2019).
113. Joo, J., Bucy, E. P. & Seidel, C. Automated coding of televised leader displays: detecting nonverbal political behavior with computer vision and deep learning. *Int. J. Commun.* **13**, 4044–4066 (2019).
114. Metallinou, A. et al. The USC CreativeIT database of multimodal dyadic interactions: from speech and full body motion capture to continuous emotional annotations. *Lang. Res. Eval.* **50**, 497–521 (2016).
115. Pouw, W., Paxton, A., Harrison, S. J. & Dixon, J. A. Acoustic information about upper limb movement in voicing. *Proc. Natl Acad. Sci. USA* **117**, 11364–11367 (2020).
116. Enfield, N., Levinson, S. C., De Ruiter, J. P. & Stivers, T. in *Field Manual* Vol. 10, 96–99 (ed. Majid, A.) (Max Planck Institute for Psycholinguistics, 2007).
117. Enfield, N. & Sidnell, J. On the concept of action in the study of interaction. *Discourse Stud.* **19**, 515–535 (2017).
118. Duran, N. D., Paxton, A. & Fusaroli, R. ALIGN: analyzing linguistic interactions with generalizable techNiques — a Python library. *Psychol. Methods* **24**, 419 (2019).
119. Brennan, S. E. & Clark, H. H. Conceptual pacts and lexical choice in conversation. *J. Exp. Psychol. Learn. Mem. Cogn.* **22**, 1482–1493 (1996).
120. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
121. Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E. & Gallant, J. L. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **532**, 453–458 (2016).
122. Shain, C. et al. Robust effects of working memory demand during naturalistic language comprehension in language-selective cortex. *J. Neurosci.* **42**, 7412–7430 (2022).
123. Fedorenko, E., Blank, I. A., Siegelman, M. & Mineroff, Z. Lack of selectivity for syntax relative to word meanings throughout the language network. *Cognition* **203**, 104348 (2020).
124. Stephens, G. J., Silbert, L. J. & Hasson, U. Speaker–listener neural coupling underlies successful communication. *Proc. Natl Acad. Sci. USA* **107**, 14425–14430 (2010).
125. Schilbach, L. et al. Toward a second-person neuroscience. *Behav. Brain Sci.* **36**, 393–414 (2013).
126. Redcay, E. & Schilbach, L. Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nat. Rev. Neurosci.* **20**, 495–505 (2019).
127. Riley, M. A., Richardson, M., Shockley, K. & Ramenzoni, V. C. Interpersonal synergies. *Front. Psychol.* **2**, 38 (2011).
128. Dale, R., Fusaroli, R., Duran, N. D. & Richardson, D. C. in *Psychology of Learning and Motivation* (ed. Ross, B. H.) Vol. 59, 43–95 (Elsevier, 2013).
129. Fusaroli, R., Rączaszek-Leonardi, J. & Tylén, K. Dialog as interpersonal synergy. *N. Ideas Psychol.* **32**, 147–157 (2014).

130. Hadley, L. V., Naylor, G. & Hamilton, A. F. D. C. A review of theories and methods in the science of face-to-face social interaction. *Nat. Rev. Psychol.* **1**, 42–54 (2022).
131. Cornejo, C., Cuadros, Z., Morales, R. & Paredes, J. Interpersonal coordination: methods achievements and challenges. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2017.01685 (2017).
132. Smaldino, P. E. in *Computational Social Psychology* 311–331 (Routledge, 2017).
133. Devezer, B., Nardin, L. G., Baumgaertner, B. & Buzbas, E. O. Scientific discovery in a model-centric framework: reproducibility, innovation, and epistemic diversity. *PLoS One* **14**, e0216125 (2019).
134. Sulik, J., Bahrami, B. & Deroy, O. The diversity gap: when diversity matters for knowledge. *Perspect. Psychol. Sci.* **17**, 752–767 (2022).
135. O'Connor, C. & Bruner, J. Dynamics and diversity in epistemic communities. *Erkenntnis* **84**, 101–119 (2019).
136. Longino, H. in *The Stanford Encyclopedia of Philosophy* (ed. Zalta, E. N.) https://plato.stanford.edu/archives/sum2019/entries/scientific-knowledge-social/ (Stanford Univ., 2019).
137. Van Rooij, I. The tractable cognition thesis. *Cogn. Sci.* **32**, 939–984 (2008).
138. Kwisthout, J., Wareham, T. & Van Rooij, I. Bayesian intractability is not an ailment that approximation can cure. *Cogn. Sci.* **35**, 779–784 (2011).
139. Contreras Kallens, P. & Dale, R. Exploratory mapping of theoretical landscapes through word use in abstracts. *Scientometrics* **116**, 1641–1674 (2018).
140. Methods for methods' sake. *Nat. Methods* https://doi.org/10.1038/nmeth1004-1 (2004).
141. Oberauer, K. & Lewandowsky, S. Addressing the theory crisis in psychology. *Psychon. Bull. Rev.* **26**, 1596–1618 (2019).
142. Meehl, P. E. Theory-testing in psychology and physics: a methodological paradox. *Phil. Sci.* **34**, 103–115 (1967).
143. Klein, O. et al. A practical guide for transparency in psychological science. *Collabra Psychol.* **4**, 20 (2018).
144. Muthukrishna, M. & Henrich, J. A problem in theory. *Nat. Hum. Behav.* **3**, 221–229 (2019).
145. Eronen, M. I. & Bringmann, L. F. The theory crisis in psychology: how to move forward. *Perspect. Psychol. Sci.* https://doi.org/10.1177/1745691620970586 (2021).
146. Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A. & Haig, B. D. Theory construction methodology: a practical framework for building theories in psychology. *Perspect. Psychol. Sci.* https://doi.org/10.1177/1745691620969647 (2021).
147. Kyvik, S. & Reymert, I. Research collaboration in groups and networks: differences across academic fields. *Scientometrics* **113**, 951–967 (2017).
148. Tebes, J. K. & Thai, N. D. Interdisciplinary team science and the public: steps toward a participatory team science. *Am. Psychol.* **73**, 549 (2018).
149. Falk-Krzesinski, H. J. et al. Mapping a research agenda for the science of team science. *Res. Eval.* **20**, 145–158 (2011).
150. da Silva, J. A. T. The Matthew effect impacts science and academic publishing by preferentially amplifying citations, metrics and status. *Scientometrics* **126**, 5373–5377 (2021).
151. Scheel, A. M., Tiokhin, L., Isager, P. M. & Lakens, D. Why hypothesis testers should spend less time testing hypotheses. *Perspect. Psychol. Sci.* **16**, 744–755 (2020).
152. Jones, M. N. *Big Data in Cognitive Science* (Psychology Press, 2016).
153. Paxton, A. & Griffiths, T. L. Finding the traces of behavioral and cognitive processes in big data and naturally occurring datasets. *Behav. Res. Methods* **49**, 1630–1638 (2017).
154. Lupyan, G. & Goldstone, R. L. Beyond the lab: using big data to discover principles of cognition. *Behav Res. Methods*, **51**, 1554–3528 (2019).
155. Haspelmath, M., Dryer, M. S., Gil, D. & Comrie, B. (eds) *The World Atlas of Language Structures* (Max Planck Digital Library, 2013).
156. Eberhard, D. M., Simons, G. F. & Fennig, C. D. *Ethnologue: Languages of the World* (SIL International, 2021).
157. MacCorquodale, K. & Meehl, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.* **55**, 95–107 (1948).
158. van Rooij, I. & Baggio, G. Theory before the test: how to build high-verisimilitude explanatory theories in psychological science. *Perspect. Psychol. Sci*. https://doi.org/10.1177/1745691620970604 (2021).
159. Christiansen, M. H. & Chater, N. *Creating Language: Integrating Evolution, Acquisition, and Processing* (MIT Press, 2016).
160. Berwick, R. C. & Chomsky, N. *Why Only Us: Language and Evolution* (MIT Press, 2016).
161. Gopnik, A. Scientific thinking in young children: theoretical advances, empirical research, and policy implications. *Science* **337**, 1623–1627 (2012).
162. Pereira, A. F., James, K. H., Jones, S. S. & Smith, L. B. Early biases and developmental changes in self-generated object views. *J. Vis.* **10**, 22–22 (2010).
163. Fagan, M. K. & Iverson, J. M. The influence of mouthing on infant vocalization. *Infancy* **11**, 191–202 (2007).
164. Martin, J., Ruthven, M., Boubertakh, R. & Miquel, M. E. Realistic dynamic numerical phantom for MRI of the upper vocal tract. *J. Imaging* **6**, 86 (2020).
165. Spivey, M. J. & Dale, R. Continuous dynamics in real-time cognition. *Curr. Dir. Psychol. Sci.* **15**, 207–211 (2006).
166. *Publication Manual of the American Psychological Association* 3rd edn (American Psychological Association, 1983).
167. *Publication Manual of the American Psychological Association* 6th edn (American Psychological Association, 2010).
168. Ashby, W. R. *An Introduction to Cybernetics* (Martino, 1956).
169. de Raadt, J. D. R. Ashby's law of requisite variety: an empirical study. *Cybern. Syst.* **18**, 517–536 (1987).
170. Ward, L. M. *Dynamical Cognitive Science* (MIT Press, 2002).
171. Regier, T., Carstensen, A. & Kemp, C. Languages support efficient communication about the environment: words for snow revisited. *PLoS One* **11**, e0151138 (2016).
172. Newell, A. *Unified Theories of Cognition* (Harvard Univ. Press, 1990).
173. Rich, P., de Haan, R., Wareham, T. & van Rooij, I. in *Proc. Annual Meeting of the Cognitive Science Society* Vol. 43, 3034–3040 (eds. Fitch, T., Lamm, C., Leder, H., and Teßmar-Raible, K.) (2021).
174. Potochnik, A. & Sanches de Oliveira, G. Patterns in cognitive phenomena and pluralism of explanatory styles. *Top. Cogn. Sci.* **12**, 1306–1320 (2020).
175. Leydesdorff, L. & Schank, T. Dynamic animations of journal maps: indicators of structural changes and interdisciplinary developments. *J. Am. Soc. Inf. Sci. Technol.* **59**, 1810–1818 (2008).
176. Leydesdorff, L. & Goldstone, R. L. Interdisciplinarity at the journal and specialty level: the changing knowledge bases of the journal *Cognitive Science*. *J. Assoc. Inf. Sci. Technol.* **65**, 164–177 (2014).
177. DeStefano, I., Oey, L. A., Brockbank, E. & Vul, E. Integration by parts: collaboration and topic structure in the CogSci community. *Top. Cogn. Sci.* **13**, 399–413 (2021).
178. Cummins, R. in *Explanation And Cognition* (eds Keil, F. C. & Wilson, R.) 117–144 (MIT Press, 2000).
179. Boyd, N. M. & Bogen, J. in *Stanford Encyclopedia of Philosophy* (ed. Zalta, E. N.) https://plato.stanford.edu/archives/win2021/entries/science-theory-observation/ (Stanford Univ., 2021).
180. Smaldino, P. E. How to build a strong theoretical foundation. *Psychol. Inq.* **31**, 297–301 (2020).
181. Chang, H. *Inventing Temperature: Measurement and Scientific Progress* (Oxford Univ. Press, 2004).

Acknowledgements

A.S.W. was supported by the National Science Foundation (grants 1529127 and 1539129/1827744) and by the James S. McDonnell Foundation (https://doi.org/10.37717/220020507). K.L.J. was supported by the National Science Foundation (grant 2017245).

Author contributions

R.D. discussed the submission theme with the editor and wrote the first draft. A.S.W. and K.L.J. refined and added to this plan and contributed major new sections of writing and revision. All authors contributed to developing the figures.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to Rick Dale.

Peer review information *Nature Reviews Psychology* thanks Berna Devezer; Michael Frank, who co-reviewed with Anjie Cao; and Justin Sulik for their contribution to the peer review of this work.