Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

Emergence of relational reasoning

Keith J Holyoak[1,2] and Hongjing Lu[1,3]

We review recent theoretical and empirical work on the emergence of relational reasoning, drawing connections among the fields of comparative psychology, developmental psychology, cognitive neuroscience, cognitive science, and machine learning. Relational learning appears to involve multiple systems: a suite of Early Systems that are available to human infants and are shared to some extent with nonhuman animals; and a Late System that emerges in humans only, at approximately age three years. The Late System supports reasoning with explicit role-governed relations, and is closely tied to the functions of a frontoparietal network in the human brain. Recent work in cognitive science and machine learning suggests that humans (and perhaps machines) may acquire abstract relations from nonrelational inputs by means of processes that enable re-representation.

Addresses
[1] Department of Psychology, University of California, Los Angeles, United States
[2] Brain Research Institute, University of California, Los Angeles, United States
[3] Department of Statistics, University of California, Los Angeles, United States

Corresponding author: Holyoak, Keith J (holyoak@lifesci.ucla.edu)

Current Opinion in Behavioral Sciences 2020, **37**:118–124

This review comes from a themed issue on **Same-different conceptualization**

Edited by **Edward A Wasserman**, **Jean-Rémy Hochmann** and **Susan Carey**

https://doi.org/10.1016/j.cobeha.2020.11.012

Introduction

Humans — and no other animals — deliberately start fires and manufacture wheels, make rules and punish those who transgress them, argue about what is true or false, construct and decipher poems and equations, predict future events based on unobservable causes, and imagine worlds that will never exist. What makes human thinking so special? A longstanding proposal [1] is that a core capacity is a high-level form of *relational reasoning* closely linked to the functions of the prefrontal cortex — what in this paper we refer to as the Late System, specific to humans. This type of reasoning has two basic prerequisites: the ability to form explicit representations of relations between entities, and the ability to make inferences by integrating multiple relations and comparing relations across domains. That is, the human brain enables a thinker to consider two or more relations together to assess what they jointly imply. By enabling inferences based on relations between entities, rather than solely on the entities themselves, the ability to generalize knowledge to new situations is dramatically increased.

Here we provide an overview of recent work that bears on the emergence of relational reasoning — how explicit relations might be acquired, and how the reasoning system available to human adults relates to cognitive capabilities exhibited by very young children and members of other species. We will review several threads of work that bear on current controversies, with a focus on analogical reasoning. These threads include behavioral and neural investigations of relational reasoning; theoretical efforts to distinguish relational reasoning from allied forms of cognition in infants and in non-human animals; and computational models of relation learning and relational reasoning, developed both in cognitive science and in machine learning. Our aim is to sketch the path by which relational reasoning appears to emerge in humans, and perhaps eventually will emerge machines. In doing so we will freely express our own opinions as to which theoretical directions appear most promising given what we currently know.

Two views of how relational reasoning emerges

In this paper we use ‘relational reasoning’ as shorthand for reasoning with higher-order, role-governed relations [2] in a manner that approximates the capabilities of a physical symbol system [3]. A relation holds between multiple entities, each of which fills a specific role (e.g. a spark has the role of cause, with fire as its effect). For adult humans, this cognitive system contributes to reasoning by analogy, making inferences based on cause and effect, evaluating moral issues, understanding the linkage between intention and action, reasoning about time and space, deductive inference, and similar manifestations of high-level cognition. (Note we say ‘contributes to’ — as we will see, additional systems also subserve high-level cognition.) For general discussions of relational reasoning see Refs. [4–6].

Perennial controversy surrounds the phylogenetic and developmental origins of the system that supports relational reasoning in adult humans. Proponents can be found for the most extreme views (no such system exists; it is ubiquitous across species from primates to birds). Here, we consider two general proposals (Box 1). One possibility is that relational reasoning is fundamentally a unitary system, probably shared to some extent with other primates (at least), that for humans emerges in early infancy. The core of the system for processing relations (perhaps some mechanism for structure mapping based on formal properties of relations; see Ref. [4]) is probably innate, although the system's complexity increases gradually to finally attain its adult level.

An alternative possibility is that in humans multiple systems support inferential processes involving relations [2]. These include several Early Systems that emerge in early infancy, which are probably shared to some extent with nonhuman animals. Each of the Early Systems is tied to a broad but constrained cognitive domain, such as causation, time, or mental attribution. None support higher-level relational reasoning, but each can produce behaviors that mimic it to some extent. Domain-general relational reasoning is a Late System, specific to humans, which comes online at about age three and develops through adolescence, coexisting throughout with the Early Systems. The Late System is a human cognitive autapomorphy [7], not shared with any other extant species. The great specialization of human cognitive evolution is a module for generalized intelligence.

Box 1 Two views of how relational reasoning emerges in humans

1 Unitary system:

- A unitary system for relational reasoning (probably shared to some extent with non-human animals) emerges in human infancy, and gradually develops into the adult system.
- The essence of the adult system (perhaps structure mapping) is present from the outset (hence presumably innate).

2 Multiple systems:

- Several Early Systems, each tied to a broad but constrained cognitive domain, such as causation, time, or mental attribution (and probably shared to some extent with non-human animals) emerge in human infancy.
- None of these Early Systems support higher-level relational reasoning, but each can produce behaviors that mimic it to some extent.
- A Late System for relational reasoning, domain-general and unique to humans, develops separately from approximately age three years to adolescence.
- The outputs of Early Systems may be re-represented in more abstract forms that can be used by the Late System.
- The Early Systems continue to function in adults, coexisting with the Late System.

The multiple systems approach is a natural extension of dual-process theories of reasoning [8]. In many cognitive domains, human judgments appear to be the products of (at least) two modules (by which we simply mean an identifiable subsystem that operates with a high degree of independence — not a Fodorian cognitive isolate). In general, the simpler modules operate from infancy, are shared to some degree with nonhuman animals, and continue to operate in human adults. But in humans, one (or more) additional modules emerge over the course of cognitive development, of which the most domain-general is the Late System for relational reasoning.

Considerable evidence supports the multiple systems view. A review of comparative research across several domains (including same/different judgments, analogy, causal relations, spatial relations, hierarchical relations, transitive relations, and mental attribution) revealed a strong discontinuity between the cognitive abilities of nonhuman animals and humans [2]. In general terms, the Early Systems shared with nonhuman animals are sufficient to perform tasks that can be solved by sensing environmental cues and responding to perceptual similarities, including those based on complex perceptual features such as symmetry and on statistical properties such as entropy (i.e. degrees of perceptual variability). In contrast, the Late System allows humans to re-represent [9] perceptual properties as explicit role-governed relations based on unobservable concepts such as causality, force, support, and time. These relations can be manipulated in working memory, enabling computations of higher-order relations between relations. Human reasoning is founded on perception, but ultimately transcends it.

More recent reviews have made the empirical case for multiple systems in the domains of causal perception and inference [10••], temporal reasoning [11••], and attribution of mental states [12••]. To take the case of time [11••], there is abundant evidence that many bird species are able to cache food, and then over extended time periods maintain a representation of its location while also updating a representation of its desirability. But such abilities can be explained by a timing mechanism that serves to update the bird's model of the environment after predetermined intervals, without assuming that the bird remembers the caching episode itself, or has any concept of time. The Early System for temporal updating can operate successfully only if information about changes in the environment is received in the same order in which those changes happen. Because time is not represented as an explicit concept, the animal is unable to reason about temporal relations when the information is received out of its temporal order, as is required to make true transitive inferences (e.g. B happened before C, and A happened before B, so A must have happened before C). A very similar limitation — an inability to reorder comparative relations based on their inherent meaning, separate from their overt order of presentation — has been observed in tests of transitive inference administered to human frontal patients [1].

One type of evidence that Early Systems coexist with, rather than develop into, the Late System is the phenomenon of 'simultaneous contradictory belief' [8, p. 11]. When two systems generate contradictory inferences, adults may be momentarily drawn to a conclusion generated by an Early System, even if this inference is then inhibited before making an overt decision. For example, if we flip a light switch at the very moment that by chance a nearby car alarm sounds, there is a strong tendency to momentarily consider that our action caused the alarm, even if prior knowledge quickly assures us a causal link is highly improbable. Similarly, when college students solve simple analogy problems under speed pressure, they will tend to direct their gaze to an incorrect option that exhibits superficial similarity to a term in the problem, leading them to be slow in responding, and to sometimes make errors [13]. The Early Systems are neither gone nor forgotten, but continue to influence cognition in adults.

Several common misconceptions about the Early Systems have been countered by careful reviews of empirical evidence. First, they are much more sophisticated than simple associative learning, involving specialized mechanisms for perceiving physical causes based on motion cues [10••,14], for temporal updating [11••], and for inferring agency from biological cues [12••]. Second, they are not all completely innate; rather, at least some undergo fast development based on experience [10••,15]. Third, the Early Systems in humans are not strictly isolated from the Late System for relational reasoning, in that outputs produced by Early Systems can provide inputs that are re-represented within the Late System. Nonhuman animals, however, lack this capacity for re-representation.

Distinguishing early from Late Systems in nonverbal and preverbal organisms

It has proven difficult to find compelling evidence that nonhuman animals exhibit higher-order relational reasoning. The classical maxim of comparative psychology known as Lloyd Morgan's canon — roughly, do not assume complex cognitive mechanisms are necessary to explain behaviors for which simpler mechanisms suffice — implies that to demonstrate that a behavior reflects the operation of the Late System, it is necessary to exclude the possibility that it might instead be produced by an Early System. In particular, empirical evidence that a thinker is reasoning using the Late System will only be provided by demonstrations of success in reasoning tasks that cannot be performed solely using perceptual cues (e.g. on the basis of direct similarity of the entities involved in training versus test examples) [16••]. For nonverbal organisms, such tests must not require language. In [2], seventeen nonverbal experimental tests of Late System relational reasoning were proposed, addressing abilities including analogy, transitive inference, understanding of weight, making causal interventions, intentional communication, and understanding of false beliefs. For example, one proposed experiment would adapt an animal conditioning paradigm that has been applied successfully in a study of causal learning with adult humans [17]. But more than a decade later, our (admittedly informal) literature review of work in comparative psychology has failed to identify a single study using any of the suggested experimental protocols.

Most work on the possible origins of relations in nonhuman animals has used variants of the array match-to-sample task (AMTS; also known as relational match-to-sample, RMTS): which is more similar to the sample AA — option BB or option CD? Substituting visual icons for letters, and given suitable reinforcement training (sometimes extreme 'dogged training', such as roughly 25 000 trials for baboons [18]), some individuals of various animal species can learn to choose the 'same' response (or the 'different' one) with accuracy above chance. Such performance has sometimes been characterized as a demonstration that nonhuman animals are able to form the abstract concepts of *same* and *different* [19•]. However, comparative psychologists have noted an alternative explanation based on perceptual entropy: the objects in the pairs AA and in BB each exhibit no variability, whereas the objects in CD are variable (and increasingly so when the displays are constructed from more than two items, all different from each other) [20]. Learning to perform AMTS can thus potentially be explained by sensitivity to a perceptual property of multi-item displays, without postulating an ability to represent an abstract relation.

Work in developmental psychology [21•] has shown that three-year-olds and four-year-olds fail 2-item AMTS (as do nonhuman animals in most studies), and that three-year-olds succeed with larger arrays only when they are able to encode stimuli in terms of entropy. By age four, the behavior of children in this task begins to contrast with that of non-human species, with robust success on 2-item AMTS achieved by age six. One infant study found that (under limited conditions) three-month-olds showed reliable habituation to 'same' or else 'different' object pairs [22], but the basis for this performance is unclear (since habituation does not involve the explicit choice required in the AMTS task). Another developmental study [23] with toddlers examined AMTS with 'fused' objects (as compared to object pairs), where fusing created a single perceptual object that is either left-right symmetrical ('same') or asymmetrical ('different'). However, this manipulation pits one complex perceptual feature (entropy) against another (symmetry), and hence the findings are open to multiple interpretations [24].

Relational reasoning in the human brain

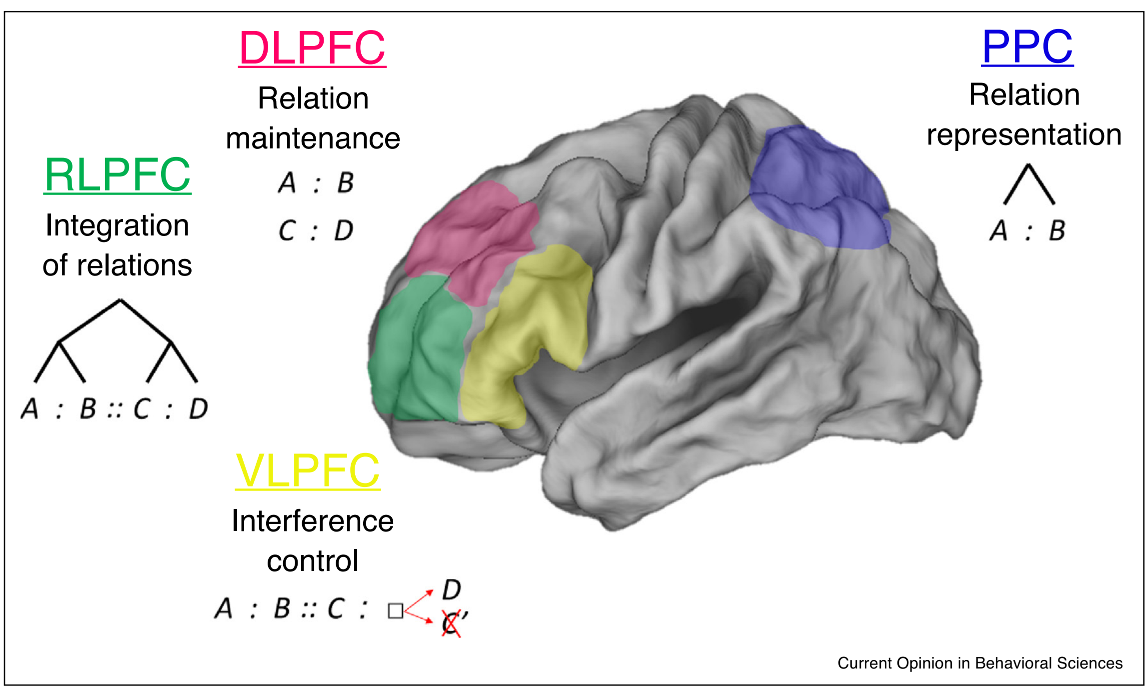
Over the past quarter century, neural investigations have identified a network of brain areas that provides the substrate for relational reasoning in humans (for reviews see Refs. [25•,26•]). Specifically, the system for relational reasoning (i.e. the Late System) depends on the neural regions sketched in Figure 1. Extensive evidence from neuroimaging, neuropsychological, and neuromodulatory studies points to a set of areas (generally left lateralized) within the frontal and parietal cortices that appear to constitute key substrates for component processes. Rostrolateral prefrontal cortex (RLPFC) plays a central role in integrating multiple relations so as to generate and/or evaluate higher-order relations (e.g. when comparing two relations to evaluate the validity of a four-term analogy, *A:B::C:D*). As one example of the abundant evidence for this conclusion, a neuropsychological study of patients with focal damage to the left RLPFC revealed specific deficits in relational integration when solving analogy problems [27]. Notably, RLPFC in the human brain differs in important ways from the homologous area in other primates [28], and the frontal cortex as a whole undergoes maturational changes over the course of childhood extending into early adulthood [29]. These evolutionary and developmental changes have implications for the emergence of relational reasoning in humans.

Other brain regions are also critical to relational reasoning (see Ref. [25•] for a review). Dorsolateral prefrontal cortex (DLPFC) is involved in maintaining relations in a working memory. Ventrolateral prefrontal cortex (VLPFC) is implicated in interference control (e.g. inhibiting salient information, often based on similarities between entities, that competes with relevant relations). Posterior parietal cortex (PPC) is implicated in the representation of individual relations (e.g. *A:B*).

Development of the Late System for relational reasoning

When we consider cognitive development in humans, based on the available evidence from multiple domains [9,10••,11••,12••], the earliest age at which we can be confident that the Late System for relational reasoning emerges is around three years. From that age forward, there is a large body of evidence concerning the developmental pattern. In general terms, as they mature children exhibit a gradual *relational shift* — an increasing focus on relations as opposed to individual objects [30]. Both culture and language have an impact in early childhood.

Figure 1



Schematic of major frontal and parietal brain areas that support representation and integration of relations, illustrated for analogical reasoning. Active representations of relations (e.g. *A:B*) are formed in posterior parietal cortex, maintained in working memory based on the DLPFC (*A:B*, *C:D*), and integrated to identify higher-order relations (e.g. sameness of the relations *A:B* and *C:D*) in rostrolateral PFC. Ventrolateral PFC is involved in inhibiting salient but potentially interfering information (e.g. *C'*, a close semantic associate of *C*, might interfere with processing the relation *C:D* that is analogous to *A:B*). Reprinted with permission from Ref. [25•].

Relational thinking seems to be accelerated in Chinese as compared to American three-year-olds [31•], and is facilitated by relevant relational vocabulary [32]. In more complex versions of RMTS using physical devices, three-year-olds attend to *same/different* relations more readily when the relations are instantiated in a manner consistent with background causal knowledge [33]. In accord with neural evidence, maturational increases in both working memory capacity and inhibitory control predict developmental progress in relational reasoning [34•,35]. Measures of fluid intelligence (associated with working memory and inhibitory control) continue to predict success in relational reasoning tasks even among college students [36,37]. For college students [38], middle-school students [39] and preschoolers [40], generating solutions to semantically-distant analogies can elicit a general relational set, facilitating subsequent processing of relations in other tasks. On a longer timescale, taking a three-month course in preparation for the Law School Admission Test (LSAT) has been found to improve performance on a dissimilar reasoning task, and also to lead to changes in brain activation measured by functional magnetic resonance imaging (fMRI) in regions associated with relational processing [41•]. Relational reasoning thus remains malleable even for highly educated adults.

The developmental pattern observed in behavioral studies of Late-System relational thinking is consistent with developmental changes in the neural system sketched in Figure 1 [42•]. Increased ability (and propensity) to focus attention on relations, greater inhibitory control, increased working memory capacity, and learning experiences that expand the repertoire of explicit relations, all support the development and effectiveness of the Late System.

There is also compelling evidence that experience generates what might be termed 'compiled relational knowledge', which supports alternative systems for high-level relational reasoning. This contrast corresponds to a shift from reliance on fluid intelligence (based on the Late System) to crystalized intelligence (based on compiled linguistic knowledge). It is notable that those frontal patients who exhibit severe deficits in relational reasoning with novel problems are sometimes relatively unimpaired in their use of language [1], which is itself an inherently relational system. Comprehension of metaphors — a type of nonliteral language — appears to draw upon multiple systems [43•,44]. Within samples of college students, comprehension of even unfamiliar metaphors (e.g. *The flowers purred in the sunlight*) can be relatively independent of measures of fluid intelligence, and is better predicted by measures of semantic knowledge [45]. In another dissociation, people with autism tend to show deficits in metaphor comprehension [46], even though analogical reasoning is an area in which they exhibit relative strength [47]. The Late System is not the sole manifestation of higher-order human thinking.

Learning relations from nonrelational inputs by computational models

Relational reasoning can potentially be achieved by nonbiological systems. Recent work in machine learning and cognitive science has been exploring algorithms that might account for more complex reasoning using relatively simple computational mechanisms. A number of machine learning models have applied forms of deep learning to extract semantic vectors (*embeddings*) representing the meanings of words, phrases, and sentences [48,49]. Such models have had considerable success in accounting for human judgments about the meanings of individual words, and efforts are underway to develop models that might deal with compositional meaning and relational inferences [50•], although so far with limited success [51–53]. Other work relies on big data and predefined knowledge bases (such as the automated thesaurus WordNet, or all possible spatial relations involved in Raven's matrix problems) to train a model to acquire relational knowledge. In [54], for example, 3.3 million word triplets (a relation label and two related entities) were used to train primarily verb-related relation types. In [55], a deep learning model was used to learn to solve Raven's matrix problems based on a training set of over a million samples (also Ref. [56]). But although these models exhibit some relational abilities after training with big data, they make the greater success of humans after training with very small data all the more impressive.

Within cognitive science, other recent work has begun to address a fundamental question previously ignored by computational models aiming to explain complex relational reasoning: How might relations be acquired from nonrelational inputs? Different avenues are being pursued [57•]; here we sketch an approach that does not assume mechanisms for structure mapping are innate. *Bayesian Analogy with Relational Transformations* (BART) emphasizes re-representation as the path of relation learning, taking semantic meanings of individual words as inputs that undergo successive transformations to acquire the representations of relations [58,59,60••]. Specifically, BART takes as inputs concatenated pairs of word embeddings created by machine learning models such as Word2vec [48]. Starting from these feature vectors for individual words, BART uses supervised statistical learning with both positive and negative examples to acquire representations of semantic relations. For example, a vector formed by concatenating the individual vectors for *old* and *young* would constitute a positive example for the *contrast* relation, and also a negative example of the *synonymy* relation.

By including a mechanism for reordering dimensions based on feature differences, the model is able to align (across word pairs) different dimensions that play comparable functional roles (e.g. dimensions that code age for *old-young* might be aligned with those that code wealth for *rich-poor*, helping the model to acquire the abstract relation *contrast*). After learning dozens of different relations individually, for any word pair BART can calculate a relation vector consisting of the posterior probability that the pair instantiates each of the learned relations, thus re-representing the raw embeddings as more meaningful distributed representations that code specific relations between individual words. These acquired relation vectors enable the model to solve simple verbal analogies with human-level accuracy [60••], predict human judgments of relational similarity [37], and predict similarity of neural responses to individual word pairs in brain regions associated with relational reasoning (Figure 1) [61]. BART serves as a proof-of-concept that analogical reasoning in humans, and perhaps machines, may emerge from a mechanism that re-represents nonrelational inputs to form explicit relations that can be manipulated and compared.

Conflict of interest statement

Nothing declared.

Acknowledgement

This research was supported by NSF Grant BCS-1827374. We thank Susan Carey for helpful comments on an earlier draft.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Waltz JA, Knowlton BJ, Holyoak KJ, Boone KB, Mishkin FS, de Menezes Santos M, Thomas CR, Miller BL: **A system for relational reasoning in human prefrontal cortex**. *Psychol Sci* 1999, **10**:119-125.

2. Penn DC, Holyoak KJ, Povinelli DJ: **Darwin's mistake: explaining the discontinuity between human and nonhuman minds**. *Behav Brain Sci* 2008, **31**:109.

3. Newell A: **Physical symbol systems**. *Cogn Sci* 1980, **4**:135-183.

4. Gentner D: **Bootstrapping the mind: analogical processes and symbol systems**. *Cogn Sci* 2010, **34**:752-775.

5. Halford GS, Wilson WH, Phillips S: **Relational knowledge: the foundation of higher cognition**. *Trends Cogn Sci* 2010, **14**:497-505.

6. Holyoak KJ: **Analogy and relational reasoning**. In *Oxford Handbook of Thinking and Reasoning.* Edited by Holyoak KJ, Morrison RG. Oxford University Press; 2012.

7. Suddendorf T: **Explaining human cognitive autapomorphies**. *Behav Brain Sci* 2008, **31**:147-148.

8. Sloman SA: **The empirical case for two systems of reasoning**. *Psych Bull* 1996, **119**:3-22.

9. Karmiloff-Smith A: *Beyond Modularity: A Developmental Perspective on Cognitive Science*. MIT Press; 1992.

10. •• Danks D: **The psychology of causal perception and reasoning**. In *Oxford Handbook of Causation.* Edited by Beebee H, Hitchcock C, Menzies P. Oxford University Press; 2009

Reviews evidence for distinct systems for causal perception, causal learning, and causal reasoning.

11. •• Hoerl C, McCormack T: **Thinking in and about time: a dual systems perspective on temporal cognition**. *Behav Brain Sci* 2019, **42**:e244

Reviews what is known about temporal processing in nonhuman animals and human infants, making the case for multiple systems.

12. •• Burge T: **Do infants and nonhuman animals attribute mental states?** *Psychol Rev* 2018, **125**:409-434

Reviews what is known about mental attribution in nonhuman animals and human infants, making the case for multiple systems.

13. Wong EF, Schauer GF, Gordon PC, Holyoak KJ: **Semantic and visual interference in solving pictorial analogies**. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society* 2019:3115-3120.

14. Kominsky JF, Scholl BJ: **Retinotopic adaptation reveals distinct categories of causal perception.** *Cognition* 2020, **203** http://dx.doi.org/10.1016/j.cognition.2020.104339.

15. Barrett HC: **A hierarchical model of the evolution of human brain specializations**. *Proc Natl Acad Sci U S A* 2012, **109**:10733-10740.

16. •• Povinelli DJ, Henley T: **More rope tricks reveal why more task variants will never lead to strong inferences about higher-order causal reasoning in chimpanzees**. *Anim Behav Cogn* 2020, **7**:392-418

Provides an analysis of the prerequisites for an experimental protocol that could demonstrate higher-level relational reasoning in nonverbal animals.

17. Shanks DR, Darby RJ: **Feature- and rule-based generalization in human associative learning**. *J Exp Psychol: Anim Behav Process* 1998, **24**:405-415.

18. Fagot J, Thompson RKR: **Generalized relational matching by guinea baboons (*Papio papio*) in two-by-two-item analogy problems**. *Psychol Sci* 2011, **22**:1304-1309.

19. • Wasserman EA, Castro L, Fagot J: **Relational thinking in animals and humans: From percepts to concepts**. In *APA Handbook of Comparative Psychology: Perception, Learning, and Cognition.* Edited by Call J, Burghardt GM, Pepperberg IM, Snowdon CT, Zentall T. American Psychological Association; 2017

A review of recent work with nonhuman animals focusing on variants of the AMTS paradigm.

20. Fagot J, Wasserman EA, Young ME: **Discriminating the relation between relations: the role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*)**. *J Exp Psychol: Anim Behav Process* 2001, **27**:316-328.

21. • Hochmann J-R, Tuerk AS, Sanborn S, Zhu R, Long R, Dempster M, Carey S: **Children's representation of abstract relations in relational/array match-to-sample tasks**. *Cogn Psychol* 2017, **99**:17-43

An extensive examination of performance on AMTS tasks with varying array sizes for children aged three-six years.

22. Anderson EM, Chang Y-J, Hesbos S, Gentner D: **Comparison within pairs promotes analogical abstraction in three-month olds**. *Cognition* 2018, **176**:74-86.

23. Walker CM, Gopnik A: **Discriminating relational and perceptual judgments: evidence from human toddlers**. *Cognition* 2017, **166**:23-27.

24. Glorioso GC, Kuznar SL, Pavlic M, Povinelli DJ: **Still no solution to non-verbal measures of analogical reasoning: reply to Walker and Gopnik (2017)**. *Cognition* 2020 http://dx.doi.org/10.1016/j.cognition.2020.104288.

25. • Holyoak KJ, Monti MM: **Relational integration in the human brain: a review and synthesis**. *J Cogn Neurosci* 2020 http://dx.doi.org/10.1162/jocn_a_01619

A recent review of evidence concerning the neural substrate of human relational reasoning.

26. • Hobeika L, Diard-Detoeuf C, Garcin B, Levy R, Volle E: **General and specialized brain correlates for analogical reasoning: a meta-analysis of functional imaging studies**. *Hum Brain Mapp* 2016, **37**:1953-1969

Summary of evidence from functional imaging studies concerning the neural substrate for analogical reasoning.

27. Urbanski M, Bréchemier M-L, Garcin B, Bendetowicz D, de Schotten MT, Foulon C, Rosso C, Clarençon F, Dupont S, Pradat-Diehl P *et al.*: **Reasoning by analogy requires the left frontal**

pole: lesion-deficit mapping and clinical implications. *Brain* 2016, **139**:1783-1799.

28. Sememdefi K, Armstrong E, Schleicher A, Zilles K, Van Hoesen GW: **Prefrontal cortex in humans and apes: a comparative study of area 10**. *Am J Phys Anthropol* 2001, **114**:224-241.

29. Caballero A, Granberg R, Tseng KY: **Mechanisms contributing to prefrontal cortex maturation during adolescence**. *Neurosci Biobehav Rev* 2016, **70**:4-12.

30. Gentner D, Rattermann MJ: **Language and the career of similarity**. In *Perspectives on Language and Thought: Interrelations in Development.* Edited by Gelman SA, Byrnes JP. Cambridge University Press; 1991.

31. • Carstensen A, Zhang J, Heyman GD, Fu G, Lee K, Walker CM: **Context shapes early diversity in abstract thought**. *Proc Nat Acad Sci U S A* 2019, **116**:13891-13896
A study that provides evidence for the influence of culture and/or language on early relational reasoning in children.

32. Simms NK, Gentner D: **Finding the *middle*: spatial language and spatial reasoning**. *Cogn Dev* 2019, **50**:177-194.

33. Walker CA, Rett A, Bonawitz E: **Design drives discovery in causal learning**. *Psychol Sci* 2020, **31**:129-138.

34. • Simms NK, Frausel R, Richland LE: **Working memory predicts children's analogical reasoning**. *J Exp Child Psychol* 2018, **166**:160-177
A study showing that individual differences in children's working memory best predicted performance on an analogical mapping task, even when controlling for age, suggesting a fundamental interrelationship between analogical reasoning and working memory development.

35. Doumas LAA, Morrison RG, Richland LE: **Individual differences in relational learning and analogical reasoning: a computational model of longitudinal change**. *Front Psychol* 2018, **24**:1235.

36. Gray ME, Holyoak KJ: **Individual differences in relational reasoning**. *Mem Cogn* 2020, **48**:96-110.

37. Ichien N, Lu H, Holyoak KJ: **Distributed code for relations predicts human judgments of relational similarity**. *J Exp Psychol: Learn Mem Cogn* 2021. in press.

38. Vendetti M, Wu A, Holyoak KJ: **Far out thinking: generating solutions to distant analogies promotes relational thinking**. *Psychol Sci* 2014, **25**:928-933.

39. DeWolf M, Bassok M, Holyoak KJ: **A set for relational reasoning: facilitation of algebraic modeling by a fraction task**. *J Exp Child Psychol* 2016, **152**:351-366.

40. Simms NK, Richland LE: **Generating relations elicits a relational mindset in children**. *Cogn Sci* 2019, **43**:e12795.

41. • Bunge SA, Leib ER: **How does education hone reasoning ability?** *Curr Dir Psychol Sci* 2020, **29**:167-173
A concise review of research on the impact of educational interventions on relational reasoning and its neural substrate.

42. • Whitaker KJ, Vendetti MS, Wendelken C, Bunge SA: **Neuroscientific insights into the development of analogical reasoning**. *Dev Sci* 2017:e12531
A study of neural changes that accompany development of analogical reasoning.

43. • Holyoak KJ, Stamenković D: **Metaphor comprehension: a critical review of theories and evidence**. *Psychol Bull* 2018, **144**:641-671
A general review of metaphor comprehension, assessing the roles played by analogical reasoning and by more language-based processes.

44. Holyoak KJ: *The Spider's Thread: Metaphor in Mind, Brain, and Poetry*. MIT Press; 2019.

45. Stamenković D, Ichien N, Holyoak KJ: **Metaphor comprehension: an individual-differences approach**. *J Mem Lang* 2019, **105**:108-118.

46. Morsanyi K, Stamenković D, Holyoak KJ: **Metaphor processing in autism: a systematic review and meta-analysis**. *Dev Rev* 2020, **57**:100925 http://dx.doi.org/10.1016/j.dr.2020.100925.

47. Morsanyi K, Stamenković D, Holyoak KJ: **Analogical reasoning in autism: a systematic review and meta-analysis**. In *Thinking, Reasoning, and Decision Making in Autism.* Edited by Morsanyi K, Byrne RMJ. Routledge; 2020.

48. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J: **Distributed representations of words and phrases and their compositionality**. *Adv Neural Inf Process Syst* 2013, **263**:111-3119.

49. Devlin J, Chang M-W, Lee K, Toutanova K: **BERT: Pretraining of deep bidirectional transformers for language understanding**. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* 2019:4171-4186.

50. • Garnelo M, Shanahan M: **Reconciling deep learning with symbolic artificial intelligence: representing objects and relations**. *Curr Opin Behav Sci* 2019, **29**:17-23
A review of recent work on relational processing from the perspective of machine learning.

51. Linzen T: **Issues in evaluating semantic spaces using word analogies**. *Proceedings of the 1st Workshop on Evaluating Vector-Space Representations for NLP* 2016:13-18.

52. Ettinger A, Elgohary A, Phillips C, Resnik P: **Assessing composition in sentence vector representations**. *Proceedings of the 27th International Conference on Computational Linguistics* 2018:1790-1801.

53. Ettinger A: **What BERT is not: lessons from a new suite of psycholinguistic diagnostics of language models**. *Trans Assoc Comput Linguist* 2020, **8**:34-48.

54. Bordes A, Glorot X, Weston J, Bengio Y: **A semantic matching energy function for learning with multi-relational data**. *Mach Learn* 2014, **94**:233-259.

55. Zhang C, Jia B, Gao F, Zhu Y, Lu H, Zhu S-C: **Learning perceptual inference by contrasting**. *Adv Neural Inf Process Syst* 2019, **32**:1073-1085.

56. Santoro A, Hill F, Barrett D, Morcos A, Lillicrap T: **Measuring abstract reasoning in neural networks**. *Proceedings of the 35th International Conference on Machine Learning* 2018:4477-4486.

57. • Martin AE, Doumas LAA: **Predicate learning in neural systems: using oscillations to discover neural structure**. *Curr Opin Behav Sci* 2019, **29**:77-83
A review of recent work on relation learning within a neural model of relational reasoning.

58. Lu H, Chen D, Holyoak KJ: **Bayesian analogy with relational transformations**. *Psychol Rev* 2012, **119**:617-648.

59. Chen D, Lu H, Holyoak KJ: **Generative inferences based on learned relations**. *Cogn Sci* 2017, **41**:1062-1092.

60. •• Lu H, Wu YN, Holyoak KJ: **Emergence of analogy from relation learning**. *Proc Nat Acad Sci U S A* 2019, **116**:4176-4181
The BART model of learning relations from nonrelational inputs, with tests of the model as a predictor of relation typicality and of solving simple verbal analogy problems.

61. Chiang JN, Peng Y, Lu H, Holyoak KJ, Monti MM: **Distributed code for semantic relations predicts neural similarity during analogical reasoning**. *J Cogn Neurosci* 2020 http://dx.doi.org/10.1162/jocn_a_01620.