COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 46 (2022) e13116

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13116

The Neural Correlates of Analogy Component Processes

John-Dennis Parsons,[a] Jim Davies[b]

[a] *Department of Psychology, Queen's University*
[b] *Department of Cognitive Science, Carleton University*

Received 10 September 2019; received in revised form 31 October 2021; accepted 21 January 2021

Abstract

Analogical reasoning is a core facet of higher cognition in humans. Creating analogies as we navigate the environment helps us learn. Analogies involve reframing novel encounters using knowledge of familiar, relationally similar contexts stored in memory. When an analogy links a novel encounter with a familiar context, it can aid in problem solving. Reasoning by analogy is a complex process that is mediated by multiple brain regions and mechanisms. Several advanced computational architectures have been developed to simulate how these brain processes give rise to analogical reasoning, like the "learning with inferences and schema abstraction" architecture and the Companion architecture. To obtain this power to simulate human reasoning, theses architectures assume that various computational "subprocesses" comprise analogical reasoning, such as analogical *access, mapping, inference*, and *schema induction*, consistent with the structure-mapping framework proposed decades ago. However, little is known about how these subprocesses relate to actual brain processes. While some work in neuroscience has linked analogical reasoning to regions of brain prefrontal cortex, more research is needed to investigate the wide array of specific neural hypotheses generated by the computational architectures. In the current article, we review the association between historically important computational architectures of analogy and empirical studies in neuroscience. In particular, we focus on evidence for a frontoparietal brain network underlying analogical reasoning and the degree to which brain mechanisms mirror the computational subprocesses. We also offer a general vantage on the current- and future-states of neuroscience research in this domain and provide some recommendations for future neuroimaging studies.

Keywords: Analogical reasoning; Relational reasoning; Rostrolateral prefrontal cortex; Dorsolateral prefrontal cortex; Lateral frontoparietal network; Four-term analogy; Component processes; Neural correlates

---

Correspondence should be sent to John-Dennis Parsons, Department of Psychology, Queen's University, Kingston, K7L 3N6 (Psychology building, Queens), ON, Canada. E-mail: jddparsons@gmail.com

1. Introduction

Analogical reasoning is a core cognitive ability underlying human learning and inference. Throughout much of history, analogies have been used to teach fundamental concepts to students, such as the law of conservation of mass (Stavy, 1991). Even if dissimilar on the surface, two things can be "analogous" to each other if they share deeper relational correspondences. The planetary model of atomic structure, for instance, showcases how solar systems and atoms can be considered analogs at a basic level of construal (Dunbar, 1997). Imagine a teacher says to her class, "electrons are to atomic nuclei as planets are to the stars." If a student unfamiliar with atomic particles heard this premise and considered possible mappings between atoms and solar systems, they may learn on their own that atoms and solar systems are alike because the constituent elements of the former (electrons and nucleus) are like the elements of the latter (planets and star) in that both sets of elements are bound by similar kinds of relations. Just as a planet "orbits" a star and is "less massive" than a star, so too does an electron "orbit" a nucleus and is "less massive" than a nucleus. Hence, considering parallel relationships between independent contexts can enable students to use their knowledge of one context (solar systems) to bootstrap their learning about another (atomic particles). In this light, analogy is a powerful cognitive mechanism for knowledge acquisition and has been viewed as a prized teaching tool for decades (Aubusson, 2006). But analogical reasoning also extends well beyond the classroom. The ability to rapidly formulate and apply analogies to everyday circumstances can aid us tremendously in areas like problem solving, communication, and creativity (Green, 2016; Green, Kraemer, Fugelsang, Gray, & Dunbar, 2012; 2017), and many researchers advocate that it is central to human cognition in general (Davies, 2009; Davies & Goel, 2001; Gentner & Smith, 2012; Hofstadter, 2001; Hofstadter & Sander, 2013).

Much is now known about the neurocognitive architecture underlying the ubiquitous human ability to reason analogically. Over the past decades, various architectures have been computationally implemented to simulate human analogical reasoning and integrate this facet of cognition into artificial intelligence systems (for a review, see Hall, 1989; Thagard, 1988). One key model that has greatly influenced modern conceptions of analogical reasoning was the original Structure-Mapping Engine (SME; Falkenhainer et al., 1986; 1989). Following SME, newer models were developed, such as the Analogical Constraint Mapping Engine (Holyoak & Thagard, 1989) and Analog Retrieval by Constraint Satisfaction (ARCS; Thagard, Holyoak, Nelson, & Gochfeld, 1990). A more modern model, called Learning and Inferences by Schema Abstraction (LISA), represents the culmination of this work, and aims to integrate core theoretical considerations from psychology and neuroscience to simulate how people reason by analogy in daily life (Hummel & Holyoak, 2003). Similarly, the Companion cognitive architecture is another model of human analogical reasoning that has garnered much attention (Forbus & Hinrich, 2017).

There is a common thread linking these cognitive architectures that distinguishes them from other approaches to simulating reasoning (e.g., Borst & Anderson, 2015; Laird et al., 2012). This common thread is Gentner's (1983) structure-mapping theory. One central feature of these structure-mapping architectures is that they break down analogical reasoning into a series of computational subprocesses. In LISA, these subprocesses have been called analogical *access*, *mapping*, *inference*, and *schema induction* (Hummel & Holyoak, 2003), and have remained consistent in computational architectures since the original implementation of SME (Falkenhainer et al., 1986; 1989). Similarly, in the Companion architecture, these subprocesses are jointly referred to as the "analogy stack" and were originally represented in different cognitive models, including SME (Falkenhainer et al., 1986; 1989), MAC/FAC (Forbus, Gentner, & Law, 1995), and SAGE (McLure et al., 2015). Yet, despite some minor differences between architectures, LISA and Companion are both fundamentally based on structure-mapping and have much in common.

In a "canonical" instance of analogical reasoning according to the structure-mapping framework, the progression of subprocesses occurs as follows. First, a person encounters a novel object or situation (called a target analog) that is difficult to understand or explain—the analogist lacks prior knowledge of this target. To better understand the target, the person may engage in analogical *access* to retrieve a familiar object or situation (called a source, or base, analog) from long-term memory. Next, information in the source analog is *mapped* to structurally similar information in the target analog. Once structural mappings have been established between the analogs, information about the familiar source can be leveraged to *infer* new information about the novel target. Finally, the structure of the analogy itself (the source, the target, and how the two relate to each other) is *inducted* as a schema into memory for use in future cases of reasoning.

Cognitive architectures propose robust accounts for how analogies are created and reasoned with and have traditionally been inspired by the workings of biological systems, specifically the human brain. A parallel area of interest, then, concerns how analogical reasoning actually occurs in the brain. Which regions are linked to analogy making? Which functions of analogy do these regions implement? In recent decades, research on the biological front has also made good progress. Since the emergence of neuroimaging techniques like functional magnetic resonance imaging (fMRI) and electroencephalography (EEG), researchers have launched many investigations to observe the neural correlates of people's thoughts and behaviors as they engage in analogical reasoning (Bunge, Wendelken, Badre, & Wagner, 2005; Cho et al., 2010; Christoff et al., 2001; 2003; Geake & Hansen, 2005; 2010; Green, Fugelsang, Kraemer, Shamosh, & Dunbar, 2006; 2010; Luo et al., 2003; Volle, Gilbert, Benoit, & Burgess, 2010; Wendelken, Nakhabenko, Donohue, Carter, & Bunge, 2008; Wharton et al., 2000). Neuropsychological studies investigating cognitive deficits in patients with brain damage due to illness or stroke have provided further empirical evidence that specific brain regions are critical to people's ability to create and reason with analogies (Aichelburg et al., 2016; Krawczyk et al., 2008; 2010; Morrison et al., 2004; Schmidt et al., 2012; Urbanski et al., 2016; Waltz et al., 1999). Some developmental studies have also shed light on how the capacity for reasoning is very much dependent on one's age, attention, and physical changes to their brain structure over time (Crone et al., 2009; Dumontheil, 2014; Richland, Morrison, & Holyoak, 2006; Rojkova et al., 2016; Thibaut & French, 2016; Wendelken, Ferrer, Whitaker, & Bunge, 2016; Whitaker, Vendetti, Wendelken, & Bunge, 2018; Wright, Matlen, Baym, Ferrer, & Bunge, 2008).

Together, psychology, computer science, and neuroscience have greatly improved our understanding of analogical reasoning, both in people and artificial intelligence systems.

However, knowledge gaps exist. One such gap concerns the computational subprocesses of analogy and the extent to which they map onto identifiable brain processes. Computational models have proposed that analogical reasoning occurs in a series of subprocesses. Meanwhile, neuroscience has linked analogical reasoning in general to specific brain networks. Yet, it remains unclear how brain networks and the individual subprocesses are related. Are different regions of the brain differentially recruited for each subprocess of analogy? Since computational models strive to implement a version of analogical reasoning that parallels processes in the human brain, it is important to understand whether these computational subprocesses are meaningfully reflected in patterns of neural activity as a person creates and reasons with analogies.

In this review, we address what is known so far from fMRI studies about the specific brain structures and functions involved in subprocesses of analogical reasoning. We suggest a set of challenges and pitfalls that current fMRI studies face when investigating the component processes of analogy. Moreover, we suggest some possible avenues that future fMRI studies may consider regarding task designs and analyses to better study the extent to which analogical subprocesses map onto biological brain processes. Ultimately, to the cognitive neuroscientist, we suggest that more studies investigating the component processes of analogy will be important to develop a comprehensive understanding of how this fundamental aspect of cognition unfolds in the brain. An improved vantage on the actual neural processes that comprise analogy making could directly improve studies of individual differences in analogical reasoning across people. This could enhance knowledge of how deficits in a particular analogical subprocess lead to difficulties in reasoning and learning more broadly. Meanwhile, to the computational modeler, new findings from neuroscience will be critical when evaluating cognitive architectures and testing the neural hypotheses they have generated. Our hope for this article is that, regardless of which field the reader most subscribes to, the topics reviewed here will highlight the benefits of cross-collaboration between psychology, computer science, and neuroscience and will help these cognitive sciences move toward an improved understanding of how we as humans learn about the world through analogies.

2. Analogical reasoning recruits a lateral, frontoparietal brain network

In the past two decades, neuroimaging research has shed light on the neural substrates of analogical reasoning. In a meta-analysis, Hobeika, Diard-Detoeuf, Garcin, Levy, and Volle (2016) reported a set of brain regions that were reliably recruited during analogical reasoning across empirical fMRI studies. Their findings showed that the most anterior, left lateral region of prefrontal cortex (PFC) is critical to analogical reasoning. This region is sometimes called the anterior PFC, frontopolar cortex, or rostrolateral prefrontal cortex (RLPFC) and is located in Brodmann area (BA) 10. In addition, Hobeika et al. (2016) also found that both the dorsolateral prefrontal cortex (DLPFC; BA 9/46) and the ventrolateral prefrontal cortex (VLPFC; BA 44/45/47) were engaged in analogical reasoning across many empirical studies. Outside of PFC, other brain regions were responsive to analogical reasoning too, including the inferior parietal lobe (IPL; BA 7/40) and insula (BA 13). Similar findings were also reported in a previous meta-analysis combining analogy and metaphor studies: the left lateral surface of PFC appeared to be especially important for analogical reasoning, and a broader frontoparietal network was observed in response to both analogies and metaphors (Vartanian, 2012). Thus, meta-analytic evidence indicates that a lateral frontoparietal network (LFPN) mediates analogical reasoning, in line with conclusions from earlier work (Green et al., 2006; Watson & Chatterjee, 2012).

There is good consensus about the different roles that these regions play in analogical reasoning. The **RLPFC (BA 10)** has long been established in relational integration, processes whereby simple sets of concrete relations are combined into more abstract and complex relational structures (Christoff et al., 2001; Kroger et al., 2002; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997; Waltz et al., 1999). One example of this is the finding that RLPFC can encode relational "categories" from instances of concrete objects, even when those objects are not perceptually similar (Davis, Goldwater, & Giron, 2017). The role of RLPFC as a hub for integrative processes is reinforced by lesion deficit studies that have shown significantly reduced ability to reason analogically when this part of the brain is damaged (Morrison et al., 2004; Urbanski et al., 2016). Within RLPFC, anterior and posterior regions are linked to differential functions, with the most anterior regions facilitating abstract integrative processing and episodic memory retrieval (Gilbert et al., 2006). Between RLPFC and adjacent regions, some evidence suggests that a rostrocaudal processing gradient exists within the PFC such that RLPFC, being the most anterior region, integrates the most complex, abstract information (Bader & Nee, 2018; Badre, 2008; Christoff, Keramatian, Gordon, Smith, & Mädler, 2009; Green et al., 2006; Krawczyk, McClelland, & Donovan, 2011).

The RLPFC is an influential brain region in analogical reasoning because of its role in abstract relational integrations. Yet, the left **DLPFC (BA 9/46)** is also vital because of its close connection to executive function and behavioral control. The DLPFC has long been linked to working memory (Manenti, Cotelli, Calabria, Maioli, & Miniussi, 2010; Petrides, 2000), which is important in analogical reasoning, when large relational structures must be held in active memory as they are compared (Morrison et al., 2001; Waltz, Lau, Grewal, & Holyoak, 2000). Badre and Nee (2018) suggest that RLPFC is more tuned to spatially and temporally abstract goal-oriented thinking, whereas DLPFC is better anatomically positioned to exert cognitive control over other brain regions based on immediate contextual considerations. In social contexts, for instance, the RLPFC might be important for integrating and representing social knowledge, while DLPFC is critical for influencing behavior based on that integrated knowledge (Steinbeis, Bernhardt, & Singer, 2012). One hypothesis is that RLPFC implements "schematic" control, while DLPFC implements "contextual" control, and so both regions exert control over lower systems but in different ways (Badre & Nee, 2018). Along these lines, Cho et al. (2010) showed that RLPFC and DLPFC activations during analogical reasoning corresponded to dissociable functions, namely, relational integration and cognitive control, respectively. Interestingly, higher-order reasoning in younger children, including analogical reasoning, seems to depend heavily on DLPFC function; yet, as children grow older, reasoning is largely offloaded to a circuit between RLPFC and IPL (Wendelken et al., 2016). Thus, DLPFC appears to play a fundamental role in analogical reasoning from early childhood; however, this role changes with age.

Another region that responds to analogical reasoning and one which has received less attention overall is left **VLPFC (BA 44/45/47)**. This region has an established role in controlling the way that semantic knowledge is accessed in memory (Badre & Wagner, 2007). Different parts of VLPFC perform different functions. While anterior VLPFC (BA 47) retrieves stored semantic knowledge, the posterior VLPFC (BA 45) engages in selection mechanisms to control which semantic knowledge is retrieved. To facilitate these controlled retrieval processes, the VLPFC has extensive connections to the medial temporal lobes that support declarative memory (Barredo, Öztekin, & Badre, 2015). Similarly, recent evidence suggests that VLPFC and anterior temporal lobes work together to facilitate semantic categorization of objects (Matsumoto, Soshi, Fujimaki, & Ihara, 2021). In analogical reasoning, the VLPFC has been shown to represent the semantic knowledge associated with source and target analogs. In one study, RLPFC and VLPFC exhibited relations with dissociable task demands: RLPFC responded more to relational integration demands, whereas VLPFC responded to the strength of semantic associations between words in verbal analogies (Bunge et al., 2005). Moreover, according to meta-analytic evidence, the VLPFC tends to activate more for verbal semantic analogies than for analogies with visuospatial stimuli across studies (Hobeika et al., 2016). The VLPFC is, therefore, important in analogical reasoning because of its ability to interface with semantic knowledge, and this role is especially evident when the analogy is high in semantic content, as in the case of verbal analogies.

Beyond PFC, it is well established that the **IPL (BA 7/40)** is another key brain region that responds to analogical reasoning. Like RLPFC, the IPL also contributes to relational reasoning (Wendelken, 2015) but unlike RLPFC, the IPL is thought to be responsible for spatial processing during visuospatial analogies and matrix problem tasks, rather than relational integration per se (Watson & Chatterjee, 2012). Indeed, the posterior parietal cortex is known more broadly for its link with spatial cognition and spatial relational processing (Ackerman & Courtney, 2012; Amorapanth, Widick, & Chatterjee, 2010). Reflecting the IPL's role in spatial cognition, visuospatial analogy and matrix problem tasks more specifically recruit regions of left IPL than do verbal analogies that do not rely on spatial processing (Hobeika et al., 2016). Beyond spatial cognition, the IPL might be an especially important brain region considering its ubiquitous role in large-scale brain network activity (Igelström & Graziano, 2017). Several networks, including the default-mode network (Raichle, 2015) and frontoparietal control network (Niendam et al., 2012; Ptak, 2012), have nodes located in IPL. If analogical reasoning is facilitated by network activity across multiple functionally connected regions, then IPL might be more central to analogical reasoning than once thought.

These brain regions all consistently activate in response to analogical reasoning tasks and together form a left LFPN underlying analogical reasoning ability in humans (Hobeika et al., 2016; Vartanian, 2012). Importantly, analogical reasoning is in fact rooted in a more fundamental kind of cognition called relational thinking (Holyoak, 2012). Relational thinking can be described as any cognitive process involving operations over *relationships* between symbolic representations, whether these are representations of objects, situations, or people (Halford, Wilson, & Phillips, 1998). Analogical reasoning is a prime example of relational thinking (Gentner, 2010) and is suggested to be the most fundamentally important quality of both human cognition (Penn, Holyoak, & Povinelli, 2008) and artificial intelligence

(Forbus, 2016). Critically, relational thinking is associated with activity in an LFPN containing the same brain regions that are linked to analogical reasoning; the left RLPFC, DLPFC, VLPFC, and IPL are all responsive to both relational thinking and analogical reasoning (Vendetti & Bunge, 2014). Therefore, a large portion of the neural activity observed in response to analogical reasoning emerges as a component of the greater LFPN underlying relational thinking in humans. An important aim for empirical work will be to study the network activity within LFPN as people reason analogically (e.g., Hammer et al., 2019), since investigations into larger attention and control networks may provide new insights into the functional organization of the human brain (Power & Petersen, 2013).

3. Neural correlates of analogical subprocesses

Although there is good evidence connecting analogical reasoning to brain regions within LFPN, less evidence is available regarding the individual subprocesses of analogy and how they relate to this brain network. How do *access*, *mapping*, *inference*, and *schema induction* relate to structures and functions in LFPN? To date, only a handful of neuroimaging studies have attempted to investigate this question (Krawczyk, McClelland, Donovan, Tillman, & Maguire, 2010; Qiu, Li, Chen, & Zhang, 2008; Volle et al., 2010; Wendelken et al., 2008). A common feature of these studies is that they all used various forms of the four-term analogy task. In four-term analogy tasks, analogies are presented as stimuli to subjects in a particular format, that is, “A is to B as C is to D” (or “A:B::C:D” for short). The terms A and B together represent the source analog, and C and D represent the target analog. Thus, the predicate “an electron is to an atomic nucleus as a planet is to a star” is a four-term analogy. Words and arbitrary symbols are acceptable as terms in a four-term analogy, but so are visual images, such as in the People Pieces Analogy Task, which uses cartoon characters (Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004).

In the following sections, we briefly review the findings from these four-term analogy neuroimaging experiments and connect them to the canonical subprocesses of analogical reasoning. We chose to explicitly consider analogical mapping, access, and inference since these are common components of analogical reasoning that are represented in different cognitive architectures (Forbus & Hinrich, 2017; Hummel & Holyoak, 2003). The Companion architecture does not reference schema induction, but we discuss schema induction here because it is a relevant component for LISA and was part of Genter’s original structure-mapping hypothesis (Gentner, 1983; Gick & Holyoak, 1983; Hummel & Holyoak, 2003). In the sections thereafter, we discuss challenges for these kinds of investigations and offer possible ways that future fMRI experiments may adapt their paradigms to better identify the relationship between computational subprocesses of analogy and biological mechanisms in the human brain.

3.1. Analogical mapping

Mapping is the core subprocess of analogy and is thought to be preceded by analogical access and followed by analogical inference. However, unlike access and inference, which are optional subprocesses that might or might not occur depending on the situation or context, mapping is the fundamental operation in every case of analogical reasoning; mapping makes the analogy (Hummel & Holyoak, 2003). Mapping is an expensive subprocess because it is an integrative function that discovers second-order relations between first-order predicates. This integrative function is constrained by structural, semantic, and pragmatic considerations. Furthermore, it requires multiple sets of elements and relations to be held in working memory as the mapping subprocess is applied (Holyoak & Thagard, 1989). Given that mapping is both a highly important subprocess and an expensive one, a key question is: where does mapping occur in the brain?

Mapping is a higher-order comparison that depends on relational integration, a cognitive function that has been strongly linked to the RLPFC (Krawczyk, 2012). Consistent with the established role of RLPFC in relational integration (Waltz et al., 1999), fMRI studies have consistently found strong activations in RLPFC and adjacent areas of PFC when people engage in analogical reasoning (Hobeika et al., 2016). Moreover, RLPFC is suggested to sit at the top of a relational processing hierarchy in PFC (Christoff et al., 2009; Green et al., 2006; Krawczyk et al., 2011), indicating that it is a main brain region for the purposes of relational integration and is most likely central to analogical mapping as well. But to what extent does activation in RLPFC actually represent higher-order comparisons occurring during the mapping subprocess? If RLPFC is the "mapping" region in the brain, then it should primarily respond to mapping demands during analogical reasoning as opposed to demands associated with the other subprocesses, including access, inference, and schema induction.

Yet, two neuroimaging studies demonstrate that this story may not be so simple. In one study, Krawczyk et al. (2010) had participants complete a four-term analogy task while in fMRI. In this task, the researchers separated each trial into three distinct phases. In the encoding phase, participants viewed two images on-screen that together constituted a source analog (e.g., images of a telescope and a boat). Then, after a small delay, the mapping/inference phase would begin, where a third image would be presented (e.g., a periscope) and the participant would try to imagine what kind of fourth image would be required to create an analogy. After another delay, in the response phase, a real fourth image would be shown (e.g., a submarine) and the participant would need to indicate via yes/no response whether the first and second images were analogous to the third and fourth images (e.g., "a telescope is to a boat as a periscope is to a submarine"). At each phase of the trial, the fMRI scanner measured neural responses. As a control, participants also completed matched semantic and perceptual tasks with the same phase-structure. Hence, this task design enabled the researchers to isolate brain responses related to analogical reasoning at each of three different task phases, namely, encoding, mapping/inference, and response.

This experiment found that several regions of PFC were more active during the analogy task than during the matched perceptual and semantic tasks, confirming the role of regions like RLPFC and DLPFC in analogical reasoning. However, the truly novel and striking finding was that regions of PFC were also more active during the encoding phase of analogical reasoning than during the mapping/inference phase. This "encoding" period activated a broad set of brain regions, implying that most of the cognitive load and processing demands were placed on this initial part of the trial, rather than the part where participants engaged in mapping and inference—the more essential parts of analogy.

In a second, similar experiment, Volle et al. (2010) had participants make judgments about six-term analogies (i.e., A: B: C:: D: E: F) while in fMRI, with trials again separated into separate encoding and mapping phases. In this experiment, participants would first see a source term on-screen (e.g., a collection of three arbitrary alphanumeric symbols varying in size, color, and other parameters), and after a delay, two similar target terms would be presented, and participants had to choose which of the two targets was more analogous to the source. In another version of the task designed to serve as a perceptual control, participants were instead required to choose the target that possessed perceptual characteristics matching those of the source. Hence, much like Krawczyk et al. (2010), this six-term visual letter paradigm enabled Volle et al. (2010) to investigate brain responses associated with the encoding and mapping phases of their tasks.

The findings from this experiment mirrored those found by Krawczyk et al. (2010). Brain activity in RLPFC during the analogical reasoning task was stronger than during the matched control task. Additionally, the same striking finding was found where the encoding phase of the experiment again elicited stronger brain activity than did the mapping phase. Neither study was able to offer a conclusive explanation for why brain responses in RLPFC were more prominent during the encoding phase than the mapping phase. The authors offered similar interpretations that early activation of RLPFC during the encoding phase may reflect a kind of semantic retrieval of rule-based knowledge. Yet, both studies also found that other memory-related regions of LFPN, such as VLPFC and DLPFC, were more active during the encoding phase than the mapping phase. In particular, the VLPFC has a well-established link with controlled retrieval of semantic and episodic memories (Badre & Wagner, 2007). Similarly, the DLPFC has long been known to play a role in working memory (Barbey, Koenigs, & Grafman, 2013). It is, therefore, unusual that RLPFC, rather than VLPFC or DLPFC, was suggested to activate for the purposes of semantic retrieval. Given that RLPFC is mainly believed to facilitate relational integration during analogical reasoning (Hobeika et al., 2016), the conclusion that RLPFC mainly serves to facilitate semantic retrieval as relations are encoded somewhat undermines conventional perspectives on this brain region.

Other work has since demonstrated that RLPFC activity is not just relevant for analogical reasoning; it is essential. In a lesion-deficit study, Urbanski et al. (2016) demonstrated that people who have damage to left RLPFC show clear deficits in analogical reasoning ability. Lesion-deficit studies have also shown that damage to other areas of PFC can hinder analogical reasoning as well (Aichelburg et al., 2016; Krawczyk et al., 2008; Morrison et al., 2004), yet Urbanski et al. (2016) were able to show the special importance of RLPFC. To date, fMRI studies have established that every kind of analogical reasoning recruits RLPFC (Hobeika et al., 2016), that analogy making cannot occur without intact RLPFC function (Urbanski et al., 2016), and that the function of RLPFC in analogy making is abstract relational integration (Christoff et al., 2001; Christoff et al., 2009; Krawczyk et al., 2011). Meanwhile, cognitive architectures have also advocated that mapping occurs in every case of analogical reasoning, that mapping is a central mechanism in analogy, and that mapping depends on abstract relational integration (Forbus & Henrich, 2017; Hummel & Holyoak, 2003). Thus, it appears that RLPFC is by far the leading candidate for a "mapping region" in the brain. Yet, there is still insufficient evidence to conclude that RLPFC activity actually *represents* analogical mapping. More studies regarding the neural correlates of analogy subprocesses are required to fully discern whether RLPFC can be considered the "mapping region."

3.2. Analogical access

Analogical access occurs when a source analog is *spontaneously* retrieved from long-term memory. It is a mechanism for recovering stored knowledge that is analogous to some target situation or problem at hand. Access might be considered a less critical subprocess than mapping, because analogical access is "optional"—not every instance of analogical reasoning will involve accessing memory to identify a source analog. For example, in the people pieces analogy task, participants are shown two pairs of cartoon characters on a computer screen (i.e., a source and a target analog). The whole four-term analogy is available from the start of the trial, so mapping can initiate without any memory search for a suitable source analog (Viskontas et al., 2004). But in other tasks, like the original fortress and tumor word problem analogy (Gentner, 1983), success depends on remembering a previous context and mapping it to the current problem, which constitutes analogical access. Whether access occurs or not thus depends on the task. In the literature, four-term analogies typically omit access demands, while analogies presented in long text or "story" format are more effective at evoking analogical access and the retrieval of a source analog from memory (Ichien, Lu, & Holyoak, 2020).

Analogical access and the neural responses associated with it has not yet been directly investigated in fMRI. However, access describes a spontaneous retrieval process that should be related to established memory systems in the brain. Spontaneous retrieval is mediated by a key memory region, the hippocampus (Dimsdale-Zucker, Ritchey, Ekstrom, Yonelinas, & Ranganath, 2018). In the PFC, retrieval of semantic knowledge has been linked to a combination of cortical regions within left VLPFC, including anterior, mid-, and posterior VLPFC regions (Badre & Wagner, 2007). Furthermore, the VLPFC has direct anatomical connections running to inferior and lateral temporal regions (Croxson et al., 2005; Petrides & Pandya, 2002a; 2002b). These lines of evidence point to the VLPFC as one region that may help mediate analogical access.

The role of VLPFC in analogical reasoning and its specific relationship with semantic retrieval was demonstrated by Bunge et al. (2005). In this fMRI study, two words were presented onscreen for about 3 s, followed by instructions and then a second pair of words for another 3 s. Participants had to compare the two sets of words and either judge whether the word sets were analogous, or judge whether the two words in the second set only were semantically related to each other. These different kinds of trials allowed testing of brain responses related to analogical reasoning versus simpler semantic comparisons. It should be noted that both kinds of trials required semantic retrieval, but the analogy condition additionally required relational integration. Importantly, the researchers also manipulated the level of associative strength between the words. This was done to identify regions involved in controlled semantic retrieval. By varying both the kind of task and strength of semantic associations, this paradigm enabled the evaluation of separate semantic retrieval and integration demands during analogical reasoning. Analysis of the neural responses revealed that both RLPFC and VLPFC activated more in analogy trials than in semantic trials. It was also found that activations in a specific region of VLPFC (BA 45) were responsive to the associative strength between words in the task. According to Bunge et al. (2005), this points to VLPFC as a noteworthy region in analogical reasoning as it may play an important role in selectively retrieving semantic knowledge from memory.

Yet, VLPFC activity during analogical reasoning can be observed even in the absence of an access component. Krawczyk et al. (2010) and Volle et al. (2010) did not have an access component in their studies, since neither study required participants to retrieve a source analog from memory to complete any of the trials. Even so, activity in VLPFC was observed during the encoding phase of both experiments. There is evidence that this region cooperates with PFC for the purposes of memory encoding (Kelley et al., 1998), and more recent findings suggest that it cooperates directly with left VLPFC during the disambiguation of semantic information (Vitello, Warren, Devlin, & Rodd, 2014). Based on these findings, it appears that VLPFC is a common brain response during analogical reasoning.

If VLPFC is important for analogical access, why does it activate during the "encoding phase," where no access demands exist? There are multiple possibilities. One possibility is that VLPFC implements general semantic retrieval and is not specific to analogical access. Access and encoding both require retrieval, but the nature of the information that is retrieved differs between access and encoding. A second possibility for the observed VLPFC activity during encoding could be that the "encoding" phase of these tasks actually elicited a kind of cognition and brain activity that extended beyond simple perceptual encoding. In the studies by Krawczyk et al. (2010) and Volle et al. (2010), participants were presented with an initial analog and told that they would need to make an analogy in the next couple of seconds. This may have led to an active "goal-state." To fulfill their provided goal, participants may have *predicted* possible analogs that could appear in the second phase of the task. This anticipation of the next stimulus might have resembled a form of analogical reasoning all on its own. It may have corresponded to a less spontaneous and more controlled version of analogical access, where participants actively made internal guesses about possible analogs that might appear next. If this were the case, then brain activity during the "encoding phase" may not have reflected any particular "stage" of analogical reasoning, but rather many smaller instances of analogical reasoning that occurred rapidly and continuously as participants anticipated the next analogy term.

Additional studies indicate that analogical access is probably linked to memory systems in hippocampus and VLPFC as many researchers have speculated. Just as analogical reasoning is fundamentally a form of relational thinking, access itself is fundamentally a form of cross-contextual relational retrieval (Jamrozik & Gentner, 2020). Relational retrieval has been investigated in fMRI and has been shown to involve both VLPFC and hippocampus. In a study by Giovanello and Schacter (2012), participants saw a series of unrelated noun pairs. For each pair, they had to create a sentence that linked them; this promoted "encoding" of the word pair. Then, in a separate retrieval task, participants were presented with word pairs again, but this time had to indicate if both words were familiar and recognizable.

Critically, the relational complexity of the word pairs was also varied to identify the differences in brain function between normal semantic retrievals and relational retrievals. They found that in adults, DLPFC, VLPFC, and hippocampus all activated more when people recognized more relationally complex word pairs. These results lend support to the possibility that analogical access is a function of memory systems in PFC and hippocampus. Additional evidence for the role of hippocampus also comes from work on unconscious analogical reasoning, which is another example of spontaneous access (Reber, Luechinger, Boesiger, & Henke, 2014).

Analogical access is somewhat underappreciated; beyond studies of relational retrieval, it has not been investigated in fMRI. This may owe in part to ongoing theoretical conflicts about how access truly operates in the brain (Hofstater & Sander, 2013; Robbins, 2017). One crucial, repeated finding regarding analogical access is that although relational information is key for analogy making, semantic information is just as important because it more readily cues retrieval processes (Jamrozik & Gentner, 2020). When reasoning analogically, people only succeed in problem solving about 10% of the time when retrieval of a source analog is guided by relational information (Genter & Smith, 2012; Gentner, 2010). However, when the retrieval of a source analog is instead guided by semantic information, the success rate of problem solving skyrockets to 60%. Indeed, although the depth of an analogy depends on relational similarities, analogy making itself is frequently guided by surface-level semantic similarities (Forbus et al., 1995). Therefore, it is important that work in neuroscience continues to develop an understanding of analogical access using neuroimaging techniques like fMRI, since better analogical access can directly enhance problem solving.

3.3. Analogical inference

Once a source analog has been retrieved from memory and structural mappings have been established between the source and target, there is an opportunity for additional inference and learning. Inference works by using prior knowledge about the source analog to bootstrap one's learning about the target. For example, if a student encounters a visualization of an atom like hydrogen, they may notice that it resembles a solar system: a small electron orbits a larger nucleus. These relational similarities can encourage creative guesses about the nature of atoms. The student may infer that if gravity holds the planets and stars together, then it must also be gravity that holds electrons and nuclei together. Yet, this would be false, and the teacher may correct that it is electromagnetism which links atomic particles together, not gravity. In this case, although the student's inference was wrong, the analogical mappings still *enabled* inference in the first place. Hence, analogies can promote learning by combining inferences with external feedback about their correctness (Hummel & Holyoak, 2003).

While the neural correlates of analogical inference have been difficult to pinpoint, some evidence is available to support the involvement of RLPFC in this component process. Using fMRI, Wendelken et al. (2008) investigated differences in brain responses between two four-term analogy tasks that leveraged word analogies (e.g., "painter is to brush as writer is to pen"). In the comparison task, participants were presented with all four words at once, and only had to make a yes/no judgment about whether the former two terms (painter is to brush) expressed a relational match with the latter two (writer is to pen). In the completion task, participants would not be presented the fourth and final word (pen); instead, they would be expected to infer it based on the initial three words. Both tasks required analogical mapping, but only the completion task required an analogical inference component.

The study found that two regions of anterior PFC were associated with the completion task and, therefore, analogical inference: left middle frontal gyrus (BA 8) and left medial frontopolar cortex (BA 10). On the other hand, during simple comparison, bilateral RLPFC and VLPFC activation (BA 10/47) was observed. This suggests that RLPFC and VLPFC were significantly active when analogical mapping alone was necessary to solve analogy problems, but not when both mapping and inference were necessary, a peculiar result given that adding a new task demand ought to increase, not decrease, the recruitment of neural resources (e.g., Cho et al., 2010). However, it has been pointed out that the lack of a concrete baseline condition could account for this result. In fact, Krawczyk et al. (2010) later found that the mapping/inference phase of analogical reasoning does recruit RLPFC and VLPFC (BA 10/47) when compared to a baseline condition. Hence, these regions of PFC tend to be stable brain activations during analogical reasoning regardless of whether inference demands are included or not.

Similar findings were found by Qiu et al. (2008). In this study, four-term analogies using lowercase letter strings (e.g., abc: abe:: ijk: ijm) were used to investigate the component processes of analogy using event-related potential (ERP) measures. Qiu et al. (2008) used a completion paradigm, wherein participants were exposed to the source analog (abc: abe) at least 2 s before presentation of the target analog (ijk: ?), at which point participants had approximately 3 s to infer the final term of the analogy. Then, a possible response (ijm) appeared on screen, and participants had to make a yes/no judgment about whether this solution was correct or not. Qiu et al. (2008) reported that in this ERP study, more activity was seen in the left PFC for the analogy tasks compared with the baseline tasks during the phase involving analogical mapping and inference. Specifically, analogical mapping and inference were associated with a robust, negative ERP component that developed in rostral PFC (BA 10) about 900–1200 ms after the presentation of target stimuli. This time window aligned with the segment of the task, where participants had to complete the analogy by inferring the final term. Additionally, at about 2000–2500 ms after target presentation, another negative ERP component was observed in dorsal PFC (BA 9), which the authors concluded may reflect maintenance demands as participants sought to verify that their inferences were correct.

These findings from Qiu et al. (2008) complement related fMRI investigations because together these studies demonstrate the recruitment of RLPFC as well as left, anterior PFC regions during analogical inference. It appears that RLPFC activations are observed ubiquitously across analogical comparison and completion tasks (i.e., whether an inference component is involved or not), consistent with claims that relational integration during mapping constitutes the core subprocess of analogy (Hummel & Holyoak, 2003). Analogical inference recruits additional regions of medial and lateral PFC. However, the precise regions recruited by analogical inference and not mapping will require additional investigations to pinpoint due to some discrepancies in task designs among existing studies (Krawczyk et al., 2010; Wendelken et al., 2008).

Another relevant finding comes from an fMRI study on relational inference (Reber, Luechinger, Boesiger, & Henke, 2012). In this study, word pairs were presented to participants subliminally, in order to have them encoded unconsciously. Some of the word pairs shared a word in common (e.g., "red-computer," "winter-red"), while some did not (e.g., "socks-sushi," "beard-dust"). Word pairs with a common word were, therefore, related to each other. In a first step, participants unconsciously encoded the word pairs. In a second step, participants had to make judgments about the relatedness of the words that they had just seen. The researchers found that when two words were related through a common word (e.g., "computer" and "winter" through "red"), the hippocampal activity at the time of subliminal encoding increased compared to when the words were not related. But furthermore, the hippocampus also activated more during *judgments* of relatedness, which involved inference demands. Therefore, a unique insight from this study is that general inference processes leverage the hippocampus, but how exactly this brain structure contributes to inference remains shrouded. One possibility is that the hippocampus and PFC regions, such as RLPFC, coactivate during analogical reasoning; they may establish functional connectedness to implement analogical inference and mapping simultaneously (Forbus & Hinrich, 2017; Hummel & Holyoak 2003). In fact, given that the other subprocesses of analogy—access and schema induction—also rely on memory processes, the PFC–hippocampus connection could be vital link in general when modeling the brain's reasoning faculties.

3.4. Analogical schema induction

Schema induction refers to a subprocess of analogical reasoning whereby the making of an analogy can ultimately create or update "schemas" in memory. Schemas are best described as complex knowledge structures that exist in and are activated by associative networks (Ghosh & Gilboa, 2014). Research on memory schemas has focused on episodic memories, and analogies can depend on episodic memories too (Reber et al., 2014). However, there is a growing appreciation in neuroscience that episodic and semantic memory processes are interdependent (Renoult, Irish, Moscovitch, & Rugg, 2019). Episodic and semantic memories may both be latent within analogical reasoning. When an analogy is presented to a person as a stimulus, knowledge about the analogs must be retrieved from memory regardless of whether the knowledge is strictly episodic or semantic. The person can then apply their retrieved knowledge to compare the source and target. The outcome of this process (e.g., the comparison or inference) can then be associated with the retrieved memory schema itself, such that the instance of analogical reasoning is consolidated to memory and can inform future instances of reasoning. In this respect, the ability to retrieve, maintain, and update knowledge in memory schemas is an integral part of analogical reasoning.

Interestingly, some studies have manipulated how schemas are initially provided to participants. For instance, in the experiment performed by Wendelken et al. (2008), in each trial the source analog was either presented as an example (e.g., "foot is to boot") or else as a specific relation (e.g., "wear"). This meant that in some trials, participants had to extract the relational schema from the example, but in other trials, they were explicitly provided with the relational schema. The results did not show any differences in brain activity contingent on the presentation of the source analog as an example or relation. However, reaction times were slower when the source analog was presented as an example, indicating that analogical reasoning was more cognitively demanding in these trials. This is presumably because when an example is available, participants have to encode the individual terms in the analog, retrieve the relational schema from memory, and then apply that schema in their problem solving; but when the source is presented as a relation, participants have immediate access to the schema and can apply it without delay.

Visuospatial analogy paradigms (Aichelburg et al., 2016; Urbanski et al., 2016; Volle et al., 2010) have used similar conditions, where participants either had to find the relational schema embedded within the source analog ("Find" condition) or else explicitly read the schema as a verbal rule ("Apply" condition). In these studies, no significant differences in participants' performance were observed between the Find and Apply conditions. These findings contrast with the results obtained from Wendelken et al. (2008) because they suggest that participants can complete the task with similar ease regardless of how the relational schema is presented. This discrepancy in findings was suggested to be due to learning effects (Volle et al., 2010). Participants may have repeatedly augmented their relational schemas in the Find condition, meaning that their ability to use these schemas could have improved over many repeated trials. Thus, performance in the Find condition may have matched the Apply condition due to relational learning between trials. Interestingly, a region of VLPFC (BA 45) in the inferior frontal sulcus (IFS) was more highly related to Find trials than Apply trials. Aichelburg et al. (2016) suggested that IFS might be involved in rule induction when no rule is explicitly provided to subjects. One possibility is that the IFS operates alongside RLPFC to represent relational schemas. The RLPFC may represent the inducted rule from IFS at a higher level of abstraction as a relational schema, consistent with the notion that RLPFC represents relational knowledge at the highest levels of abstraction (Bader & Nee, 2018; Badre, 2008; Christoff et al., 2009; Green et al., 2006; Krawczyk et al., 2011).

An important point to note about schema induction is that it is highly related to the framework of concept learning (Zeithamova et al., 2019) and the learning of relational categories (Gentner & Kurtz, 2005). In this framework, relational schemas are thought to be accessed through a neural pathway that links ventromedial prefrontal cortex (VMPFC) to the hippocampus (Gilboa & Marlatte, 2017; Preston, Shrager, Dudukovic, & Gabrieli, 2004, Zeithamova, Dominick, & Preston, 2012). Additionally, the default-mode network is believed to play a critical role in mediating episodic memory retrieval through this neural pathway (Raichle, 2015; Smallwood et al., 2021). Several of the studies discussed here have reported activity in medial regions of PFC during their analogy tasks (Krawczyk et al., 2010; Volle et al., 2010; Wendelken et al., 2008). However, VMPFC is discussed little in the analogy literature because it is a less robust neural correlate. The brain regions underlying analogical reasoning tend to be lateral rather than medial (Hobeika et al., 2016). Furthermore, it is well established that activity in VMPFC represents the subjective and emotional *values* of things, but value is not usually manipulated in analogy tasks; this may explain why analogy-induced fluctuations in VMPFC activity are minimal (Clithero & Rangel, 2014; Winecoff et al., 2013). On the other hand, the VLPFC is a far more involved brain region in analogical reasoning, and VLPFC also has links to hippocampus (Giovanello & Schacter, 2012) and memory retrieval

(Badre & Wagner, 2007; Barredo et al., 2015; Wais, Kim, & Gazzaley, 2012). Yet, in fMRI studies of analogical reasoning to-date, it is unclear to what extent activity in VMPFC or VLPFC may have corresponded to the maintenance of relational schemas. It is also unknown to what extent default-mode network activity may have mediated maintenance of relational schemas. Hence, inducing new memory schemas during analogical reasoning may require VMPFC or VLPFC contributions, but the exact nature of these contributions and whether they depend on larger functional brain networks remains largely unexplored.

4. Isolating the analogy subprocesses in fMRI: Challenges and pitfalls

One limitation of the work reviewed above is that some subprocesses of analogical reasoning have received much more attention in fMRI studies than others. These studies have attempted to measure brain responses differentially associated with specific subprocesses of analogical reasoning, yet none has attempted to study brain responses across all four canonical subprocesses of analogy, and none has directly investigated the access, inference, and schema induction subprocesses. A very straightforward explanation for this is that the subprocesses are quite interdependent in the brain and, therefore, difficult to isolate and measure. Hummel and Holyoak (2003) note that “analog retrieval, mapping, inference, and schema induction, as well as rule-based reasoning, are all special cases of the same sequential, directional algorithm for mapping.” Mapping is the core subprocess in both the LISA and Companion cognitive architectures because mapping reflects the fundamental integrative operation that every instance of analogical reasoning relies on. This has motivated formal investigations of analogical mapping in fMRI (Krawczyk et al., 2010; Volle et al., 2010; Wendelken et al., 2008). Yet, the neural correlates of analogical mapping are still not completely understood (Urbanski et al., 2016). Furthermore, the other subprocesses have received even less attention. While analogical access and inference have been investigated in fMRI, the available evidence for these subprocesses in the brain is limited (Giovanello & Schacter, 2012; Wendelken et al., 2008).

Isolating the subprocesses is a challenge in part because different subprocesses require specialized task designs to measure, as well as reasonable subtraction tasks for fMRI. The information presented in a four-term analogy task determines which subprocesses are necessarily evoked when a person completes the task. A good example of this was shown by Wendelken et al. (2008), who manipulated brain function related to analogical inference, and to an extent schema induction, by simply adjusting the four-term analogy task to provide different information to participants. In their paradigm, presenting only three terms and instructions to infer, as opposed to presenting all four terms and instructions to compare, made the difference between “analogy with inference” and “analogy without inference.” Similarly, presenting the source analog as a “relation” or as an “example” altered the way that semantic information had to be accessed in memory and applied during the task, which also affects the nature of schema induction. Indeed, not all analogies are created equally: the kinds of information presented in an analogy can affect the kinds of brain mechanisms that are recruited when one reasons analogically (Schmidt et al., 2012).

Importantly, beyond the source and target analogs, the instruction or "goal" of the task is also a relevant pragmatic cue that can help to determine which subprocesses will initiate. If a person were shown a target analog and instructed to adopt the explicit goal of creating an analogy, this would create an "encoding" phase like the ones investigated by both Krawczyk et al. (2010) and Volle et al. (2010). On the other hand, if a person were shown a target analog but given no goal to create an analogy, there is a far smaller chance that they would be spontaneously reminded about a relevant source analog, which would constitute a case of analogical access. The presence of a "goal" can, therefore, alter cognitive demands imposed by the task, including the subprocesses that are elicited. Since analogical reasoning is a way of processing information that puts new experiences in the context of older ones, the specific *kinds* of informative cues available in the environment (including not just the source and target, but other contextual and pragmatic cues like the goal) have a noteworthy impact on the precise subprocesses that are required at each instant.

This issue of "teasing out" brain activity related to specific subprocesses of analogy is further confounded by another challenge related to neuroimaging limitations. Cognitive architectures postulate that the subprocesses of analogy are temporally overlapping processes in the brain; in other terms, they are parallel processes (Hummel & Holyoak, 2003). Indeed, all subprocesses of analogy are considered to be part of a larger "guided pattern search" mechanism that constitutes the core mapping process. This guided pattern search is both a fast process and a complex one, with distributed patterns of neural activity in many regions of LFPN contributing to analogical reasoning on a very sensitive timescale. According to these architectures, mapping occurs in only hundreds of milliseconds and is a function of dynamic spreading activation throughout associative networks. Hence, there is enormous difficulty in temporally isolating the subprocess and measuring that activity using fMRI because of the poor temporal resolution that fMRI affords.

Studies that have separated their tasks into separate encoding and mapping phases represent a good initial attempt at measuring brain activity associated with distinct subprocesses (Krawczyk et al., 2010; Volle et al., 2010). This technique continues to be a primary means of analysis (Chiang, Peng, Lu, Holyoak, & Monti, 2021). By presenting different cues at different points in time, this paradigm "slows" the analogical reasoning process enough for fMRI to obtain multiple measurements in different temporal phases. Yet, there are pitfalls that come with such a design. The first is that during the encoding phase, participants may have constantly anticipated the presentation of the fourth and final analogy term. In this respect, the encoding phase may have actually corresponded to several rapid cases of analogical inference. This ties into a second pitfall, which is that encoding is not a canonical analogical subprocess. "Encoding," which refers to the activation of associative networks in response to external cues, is not a core facet of analogy making, but rather a general perceptual process that is required in every case of analogical reasoning. The extensive LFPN activity observed in these studies during the "encoding" phase is probably not explained by perceptual encoding alone. On the other hand, it is conceivable that a form of rapid, preemptive analogy making occurs when people encode a source analog while knowing that the task will soon present a target and ask for a judgment about the validity of the analogy.

Since analogical reasoning occurs within a very small temporal window, wherein brain activations overlap and influence each other, EEG might be a promising avenue, as it provides superior temporal resolution. This was the strategy used by Qiu et al. (2008), who framed their investigation around schema induction. Yet, EEG is accompanied by the opposite problem of poor spatial resolution, in that it cannot measure precise, spatially distributed brain activations. This is perhaps an even greater limitation toward identifying the neural correlates of the analogical subprocesses. Even if the subprocesses occur in parallel and are difficult to temporally isolate, they still may have their own unique "spatial signatures" in the brain. Indeed, this seems to be the case: different kinds of four-term analogies recruit different subsets of brain regions in LFPN (Hobeika et al., 2016). Thus, when considering how analogical subprocesses relate to actual brain processes, it appears that specific cues in four-term tasks lead to cognitive demands for particular subprocesses, which is reflected by differential activation of the LFPN across regions like RLPFC, DLPFC, VLPFC, and IPL.

5. Avenues for future fMRI investigations: Task designs and analysis methods

Based on the established cognitive architectures that strongly feature analogy (LISA and Companion), there are two premises regarding the "canonical progression" of analogy subprocesses that future fMRI studies may consider when designing paradigms to isolate and measure associated activity in the brain. The first premise is that very few laboratory four-term analogy tasks elicit "canonical" cases of analogical reasoning. Different kinds of tasks elicit different subsets of analogical subprocesses; access, mapping, inference, and schema induction are not all measurable in every case of reasoning. It is possible to design four-term analogy tasks that omit access and inference demands entirely. Mapping demands can never be omitted due to the fundamental nature of the subprocess. Similarly, the extent to which schema induction demands can be selectively omitted from a task is also questionable. Schema induction relates to how knowledge is consolidated, accessed, and ultimately maintained in memory as part of a larger structure. Analogical schema induction may, therefore, have deep, interdependent ties with both analogical access and inference when examining its underlying neural components.

The second premise is that, even in cases of reasoning where all analogical subprocesses are known to play a role, it is not possible to temporally isolate these subprocesses into different task phases using fMRI. Analogical reasoning is a rapid process that occurs in only a few hundred milliseconds and is mediated by a constrained search algorithm. This algorithm simultaneously implements multiple subprocesses. The ARCS and MAC/FAC architectures, for instance, both postulate that to quickly retrieve a relevant source analog from memory, analogical access and mapping must occur simultaneously, where structural, contextual, and pragmatic factors all constrain the search for a source analog in real time. Similarly, analogical inference can be construed as an additional outcome of the search, rather than a temporally defined phase. Analogical subprocesses are not independent, temporally ordered neural mechanisms; they are better construed as different facets of one ubiquitous, underlying search process.

How, then, do future fMRI studies approach the challenge of disentangling separable neural systems for access, mapping, inference, and schema induction? Several avenues exist. The first is to develop better four-term analogy tasks that target specific component processes. In particular, more attention should be devoted to access and inference, since understanding the neural correlates of these subprocesses may shed new light on schema induction as well. One study to look toward for an example is Wendelken et al. (2008). This study differed from other noteworthy studies discussed (Krawczyk et al., 2010; Volle et al., 2010) in that it did not attempt to separate out the task into phases ordered in time. Instead, they investigated neural activity related to mapping with inference and mapping without, and they induced these alternative kinds of analogical reasoning by varying the information presented to participants. In this respect, the authors investigated the "core mapping" subprocess and its "inference facet." Yet, despite the need, analogical access has not been directly investigated in fMRI. Previous work has shown that spontaneous relational retrieval can be induced by first training participants on a set of stimuli and subsequently re-exposing them to those stimuli (Jamrozik & Gentner, 2020). This approach was used in fMRI to investigate relational retrieval (Giovanello & Schacter, 2012), relational inference (Reber et al., 2012), and unconscious analogical reasoning (Reber et al., 2014). It has also been used in fields related to analogical reasoning, including retrieval-mediated learning (Zeithamova et al., 2012) and role-governed categorization (Goldwater & Markman, 2011; Goldwater, Markman, & Stilwell, 2011).

What is now needed is a similar four-term analogy task to investigate the neural correlates of spontaneous analogical access. This could be accomplished by first training participants to recognize a set of source analogs (A:B pairs) that exemplify specific relational categories. Then, in fMRI, a new set of target analogs (C:D pairs) could be presented, and participants would indicate by button press if the target analog belongs to a familiar relational category or a novel one. Since analogical access and mapping occur simultaneously as the source analog is retrieved, measurements of brain responses at the time of button press would be expected to reveal both RLPFC (relational integration during mapping) and VLPFC (relational and semantic retrieval during access) activation if the relational category were familiar. This task should be compared to a standard analogical comparison task (i.e., mapping without access or inference) to tease out brain responses that are dedicated specifically to access.

Beyond task design, there are more advanced fMRI analysis methods that can better disentangle which brain responses are related to a specific subprocess and which brain responses are not. Specifically, multivoxel pattern analysis (MVPA; Norman, Polyn, Detre, & Haxby, 2006; Weaverdyck, Lieberman, & Parkinson, 2020) is an analysis technique that can help identify specific neural markers for analogical access, mapping, inference, and schema induction. While traditional univariate fMRI methods indicate *how much* a region activates for a specific subprocess, MVPA can better address questions about whether activity in a given region actually *represents* one subprocess or another. Although some use of multivariate fMRI is already evident in the analogy literature (e.g., Chiang et al., 2021; Westphal, Reggente, Ito, & Rissman, 2016), this approach has not yet been deployed to investigate the subprocesses of analogical reasoning, even though it has the potential to address new and important questions.

Furthermore, another analysis approach that can boost insight into the neural correlates of the subprocesses is model-based fMRI. In this approach, a mathematical model of human behavior is used to calculate a behavioral measure for participants, which can then be correlated with the BOLD signal to draw conclusions about brain regions that are functionally tuned to the behavioral measure. As discussed, several advanced computational architectures exist that simulate human analogical reasoning and predict such behavioral measures (Forbus & Hinrich, 2017; Hummel & Holyoak, 2003). Researchers have applied model-based fMRI to investigate the component stages of problem solving (Anderson & Fincham, 2014), and past neurocomputational models of analogical reasoning have proven successful too (Morrison et al., 2004). Yet, the approach has not been deployed to investigate the subprocesses of analogy. A modern neurocomputational model that integrates computational architectures in the analysis of fMRI data may show an enhanced capability to identify signals unique to different subprocesses. Such findings could enable more novel and precise insights into how the brain mediates processes like access, mapping, inference, and schema induction.

Another helpful avenue for future fMRI work would be to replicate lesion-deficit studies of analogical reasoning ability (Aichelburg et al., 2016; Morrison et al., 2004; Urbanski et al., 2016), but with specific attention paid to performance in each of the subprocesses depending on lesion deficits. Similarly, brain stimulation studies, using techniques like transcranial magnetic stimulation (TMS), could be used to temporarily perturb brain regions on the lateral surface. These techniques would allow researchers to link brain regions more conclusively to analogy subprocesses if such studies were to show that a loss of brain function in LFPN can lead to a loss of analogical reasoning performance due to some subprocesses and not others. For example, one TMS study showed that stimulating the VLPFC caused difficulties in episodic retrieval (Wais et al., 2012). Would perturbing the VLPFC during a four-term analogy task also cause drops in task performance? If so, to what extent would the performance drop reflect issues with a given subprocess, like analogical access? By leveraging diverse neuropsychological study designs, like lesion-deficit or TMS studies, it will be possible to paint a clearer picture of the brain mechanisms underlying analogy subprocesses.

Finally, although a specific network—the LFPN—has been linked to analogical and relational reasoning (Hobeika et al., 2016; Vendetti & Bunge, 2014), it is still not clear how the analogical subprocesses may depend on larger functional networks, such as the default-mode network (Raichle, 2015) or multiple demand network (Duncan, 2010). One fMRI study investigated network activity related to analogical reasoning in a large sample of subjects using MVPA (Hammer et al., 2019). They found four different networks of coactivated brain regions during the analogy task. These networks corresponded to visuospatial cognition, executive function, default-mode activation, and attention to saliency. Interestingly, activations in the default-mode network were, generally, negatively related to performance in the analogy task, whereas activations in executive frontoparietal regions were positively related. This aligns both with existing knowledge about the LFPN underlying analogy and the general understanding that the default-mode network is more important for internal mentation than focus to external tasks (Brewer et al., 2011). However, a question for future work is whether different kinds of analogy exhibit different relations with these networks depending on the subprocesses involved. A case of analogical reasoning with access may require greater activations in default-mode network to retrieve stored information in memory, compared to if no access demands were motivated by the task. In analogies that require access, it is possible that default-mode and executive-control networks cooperate to simultaneously retrieve internal knowledge and reason about analogy problems in a goal-directed fashion (Gerlach, Spreng, Gilmore, & Schacter, 2011); but this is speculation. How analogical access, mapping, inference, and schema induction separately recruit large brain networks remains to be seen and is an open area for investigation.

6. Conclusion

Computational architectures have the powerful ability to simulate human analogical reasoning and make a number of explicit predictions about how this process unfolds in the brain. Similarly, fMRI studies have made great progress toward uncovering the actual neural mechanisms that underpin analogical reasoning. However, there is less consensus surrounding how computational subprocesses of analogy relate to brain mechanisms. Here, we have provided an initial vantage on how computational architectures and fMRI studies have informed each other in studies of the analogical subprocesses. Mapping is a core subprocess that may primarily depend on intact RLPFC function in the brain. Analogical access, also called retrieval, is another important subprocess of analogical reasoning and appears linked to VLPFC and hippocampus function. Analogical inference has been linked to multiple regions, including RLPFC, VLPFC, and even hippocampus. Similarly, analogical schema induction may be a broader process that depends on similar neural mechanisms as access, inference, and potentially even mapping. The available evidence from fMRI studies, therefore, lends good credibility to the hypothesis that these subprocesses are mediated by overarching and overlapping neural mechanisms in the brain (Hummel & Holyoak, 2003).

Identifying the precise nature of these neural mechanisms is an important endeavor for future work. Analogical reasoning is often hailed as an integral aspect of human cognition. It is believed to reside at the core of the human ability to learn and apply complex relational knowledge to solve problems in the real world. Such an important process deserves to be well understood. But to understand analogical reasoning, we need to understand its component processes and how each one is implemented in the biological brain. Future work in neuroscience should strive to more stringently test the concrete hypotheses offered by computational architectures. Additionally, computational architectures may undergo continuous refinement as more is learned about how the brain reasons analogically.

Acknowledgments

Thanks to: Carleton University for project support, the Science of Imagination Laboratory for critical feedback, colleagues and peers for insight and encouragement, and all others for their own contributions.

Conflicts of interest

The authors have no conflicts to disclose.

Funding

There is no funding associated with this work.

Data availability statement

No datasets were generated or analyzed during this review.

References

Ackerman, C. M., & Courtney, S. M. (2012). Spatial relations and spatial locations are dissociated within prefrontal and parietal cortex. *Journal of Neurophysiology*, *108*(9), 2419–2429. https://doi.org/10.1152/jn.01024.2011

Aichelburg, C., Urbanski, M., Schotten, M. T. D., Humbert, F., Levy, R., & Volle, E. (2016). Morphometry of left frontal and temporal poles predicts analogical reasoning abilities. *Cerebral Cortex*, *26*(3), 915–932. https://doi.org/10.1093/cercor/bhu254

Amorapanth, P. X., Widick, P., & Chatterjee, A. (2010). The neural basis for spatial relations. *Journal of Cognitive Neuroscience*, *22*(8), 1739–1753. https://doi.org/10.1162/jocn.2009.21322

Anderson, J. R., & Fincham, J. M. (2014). Extending problem-solving procedures through reflection. *Cognitive Psychology*, *74*, 1–34. https://doi.org/10.1016/j.cogpsych.2014.06.002

Aubusson, P. (2006). *Can analogy help in science education research?*. https://doi.org/10.1007/1-4020-3830-5_14

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*(5), 193–200. https://doi.org/10.1016/j.tics.2008.02.004

Badre, D., & Nee, D. E. (2018). Frontal cortex and the hierarchical control of behavior. *Trends in Cognitive Sciences*, *22*(2), 170–188. https://doi.org/10.1016/j.tics.2017.11.005

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883–2901. https://doi.org/10.1016/j.neuropsychologia.2007.06.015

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*(5), 1195–1205. https://doi.org/10.1016/j.cortex.2012.05.022

Barredo, J., Öztekin, I., & Badre, D. (2015). Ventral fronto-temporal pathway supporting cognitive control of episodic memory retrieval. *Cerebral Cortex*, *25*(4), 1004–1019. https://doi.org/10.1093/cercor/bht291

Borst, J. P., & Anderson, J. R. (2015). Using the ACT-R cognitive architecture in combination with fMRI data. In B. U. Forstmann & E.-J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 339–352). Springer. https://doi.org/10.1007/978-1-4939-2236-9_17

Brewer, J. A., Worhunsky, P. D., Gray, J. R., Tang, Y.-Y., Weber, J., & Kober, H. (2011). Meditation experience is associated with differences in default mode network activity and connectivity. *Proceedings of the National Academy of Sciences*, *108*(50), 20254–20259. https://doi.org/10.1073/pnas.1112029108

Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex*, *15*(3), 239–249. https://doi.org/10.1093/cercor/bhh126

Chiang, J. N., Peng, Y., Lu, H., Holyoak, K. J., & Monti, M. M. (2021). Distributed code for semantic relations predicts neural similarity during analogical reasoning. *Journal of Cognitive Neuroscience*, *33*(3), 377–389. https://doi.org/10.1162/jocn_a_01620

Cho, S., Moody, T. D., Fernandino, L., Mumford, J. A., Poldrack, R. A., Cannon, T. D., Knowlton, B. J., & Holyoak, K. J. (2010). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. *Cerebral Cortex*, *20*(3), 524–533. https://doi.org/10.1093/cercor/bhp121

Christoff, K., Keramatian, K., Gordon, A. M., Smith, R., & Mädler, B. (2009). Prefrontal organization of cognitive control according to levels of abstraction. *Brain Research*, *1286*, 94–105. https://doi.org/10.1016/j.brainres.2009.05.096

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. E. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage*, *14*(5), 1136–1149. https://doi.org/10.1006/nimg.2001.0922

Christoff, K., Ream, J. M., Geddes, L. P. T., & Gabrieli, J. D. E. (2003). Evaluating self-generated information: Anterior prefrontal contributions to human cognition. *Behavioral Neuroscience*, *117*(6), 1161–1168. https://doi.org/10.1037/0735-7044.117.6.1161

Clithero, J. A., & Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*, *9*(9), 1289–1302. https://doi.org/10.1093/scan/nst106

Crone, E. A., Wendelken, C., Leijenhorst, L. V., Honomichl, R. D., Christoff, K., & Bunge, S. A. (2009). Neurocognitive development of relational reasoning. *Developmental Science*, *12*(1), 55–66. https://doi.org/10.1111/j.1467-7687.2008.00743.x

Croxson, P. L., Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Pinsk, M. A., Gross, C. G., Richter, W., Richter, M. C., Kastner, S., & Rushworth, M. F. S. (2005). Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *Journal of Neuroscience*, *25*(39), 8854–8866. https://doi.org/10.1523/JNEUROSCI.1311-05.2005

Davies, J., & Goel, A. K. (2001). Visual analogy in problem solving. In IJCAI (pp. 377–384).

Davies, J. (2009). Experience-based reasoning as the basis of a general artificial intelligence architecture. In IJCAI Workshop on Grand Challenges for Reasoning from Experiences (pp. 1–6). Pasadena, CA.

Davis, T., Goldwater, M., & Giron, J. (2017). From concrete examples to abstract relations: The rostrolateral prefrontal cortex integrates novel examples into relational categories. *Cerebral Cortex*, *27*(4), 2652–2670. https://doi.org/10.1093/cercor/bhw099

Dimsdale-Zucker, H. R., Ritchey, M., Ekstrom, A. D., Yonelinas, A. P., & Ranganath, C. (2018). CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nature Communications*, *9*(1), 294. https://doi.org/10.1038/s41467-017-02752-1

Dumontheil, I. (2014). Development of abstract thinking during childhood and adolescence: The role of rostrolateral prefrontal cortex. *Developmental Cognitive Neuroscience*, *10*, 57–76. https://doi.org/10.1016/j.dcn.2014.07.009

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 461–493). American Psychological Association. https://doi.org/10.1037/10227-017

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1986). *The structure-mapping engine*.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, *41*(1), 1–63. https://doi.org/10.1016/0004-3702(89)90077-5

Forbus, K. D. (2016). Software social organisms: Implications for measuring AI progress. *AI Magazine*, *37*(1), 85–90. https://doi.org/10.1609/aimag.v37i1.2648

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, *19*(2), 141–205. https://doi.org/10.1207/s15516709cog1902_1

Forbus, K. D., & Hinrich, T. (2017). Analogy and relational representations in the companion cognitive architecture. *AI Magazine*, *38*(4), 34–42. https://doi.org/10.1609/aimag.v38i4.2743

Geake, J. G., & Hansen, P. C. (2005). Neural correlates of intelligence as revealed by fMRI of fluid analogies. *Neuroimage*, *26*(2), 555–564. https://doi.org/10.1016/j.neuroimage.2005.01.035

Geake, J. G., & Hansen, P. C. (2010). Functional neural correlates of fluid and crystallized analogizing. *Neuroimage*, *49*(4), 3489–3497. https://doi.org/10.1016/j.neuroimage.2009.09.008

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10.1016/S0364-0213(83)80009-3

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34*(5), 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W.-k. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman & P. Wolff, *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 151–175). American Psychological Association. https://doi.org/10.1037/11156-009

Gentner, D., & Smith, L. (2012). Analogical reasoning. *Encyclopedia of Human Behavior*, *2*, 130–136.

Gerlach, K. D., Spreng, R. N., Gilmore, A. W., & Schacter, D. L. (2011). Solving future problems: Default network and executive activity associated with goal-directed mental simulations. *Neuroimage*, *55*(4), 1816–1824. https://doi.org/10.1016/j.neuroimage.2011.01.030

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. *Journal of Cognitive Neuroscience*, *18*(6), 932–948. https://doi.org/10.1162/jocn.2006.18.6.932

Gilboa, A., & Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends in Cognitive Sciences*, *21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Giovanello, K. S., & Schacter, D. L. (2012). Reduced specificity of hippocampal and posterior ventrolateral prefrontal activity during relational retrieval in normal aging. *Journal of Cognitive Neuroscience*, *24*(1), 159–170. https://doi.org/10.1162/jocn_a_00113

Goldwater, M. B., & Markman, A. B. (2011). Categorizing entities by common role. *Psychonomic Bulletin & Review*, *18*(2), 406–413. https://doi.org/10.3758/s13423-011-0058-0

Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition*, *118*(3), 359–376. https://doi.org/10.1016/j.cognition.2010.10.009

Green, A. E. (2016). Creativity, within reason: Semantic distance and dynamic state creativity in relational thinking and reasoning. *Current Directions in Psychological Science*, *25*(1), 28–35. https://doi.org/10.1177/0963721415618485

Green, A. E., Fugelsang, J. A., Kraemer, D. J. M., Shamosh, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research*, *1096*(1), 125–137. https://doi.org/10.1016/j.brainres.2006.04.024

Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: Semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex*, *20*(1), 70–76. https://doi.org/10.1093/cercor/bhp081

Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2012). Neural correlates of creativity in analogical reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2), 264–272. https://doi.org/10.1037/a0025764

Green, A. E., Spiegel, K. A., Giangrande, E. J., Weinberger, A. B., Gallagher, N. M., & Turkeltaub, P. E. (2017). Thinking cap plus thinking zap: TDCS of frontopolar cortex improves creative analogical reasoning and facilitates conscious augmentation of state creativity in verb generation. *Cerebral Cortex*, *27*(4), 2628–2639. https://doi.org/10.1093/cercor/bhw080

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*(6), 803–831. https://doi.org/10.1017/S0140525X98001769

Hall, R. P. (1989). Computational approaches to analogical reasoning: A comparative analysis. *Artificial Intelligence*, *39*(1), 39–120. https://doi.org/10.1016/0004-3702(89)90003-9

Hammer, R., Paul, E. J., Hillman, C. H., Kramer, A. F., Cohen, N. J., & Barbey, A. K. (2019). Individual differences in analogical reasoning revealed by multivariate task-based functional brain imaging. *Neuroimage*, *184*, 993–1004. https://doi.org/10.1016/j.neuroimage.2018.09.011

Hobeika, L., Diard-Detoeuf, C., Garcin, B., Levy, R., & Volle, E. (2016). General and specialized brain correlates for analogical reasoning: A meta-analysis of functional imaging studies. *Human Brain Mapping*, *37*(5), 1953–1969. https://doi.org/10.1002/hbm.23149

Hofstadter, D. R. (2001). *Analogy as the core of cognition.*

Hofstadter, D. R., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. Basic Books.

Holyoak, K. J. (2012). *Analogy and relational reasoning*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199734689.013.0013

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*(3), 295–355. https://doi.org/10.1207/s15516709cog1303_1

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*, 220–264.

Ichien, N., Lu, H., & Holyoak, K. J. (2020). Verbal analogy problem sets: An inventory of testing materials. *Behavior Research Methods*, *52*(5), 1803–1816. https://doi.org/10.3758/s13428-019-01312-3

Igelström, K. M., & Graziano, M. S. A. (2017). The inferior parietal lobule and temporoparietal junction: A network perspective. *Neuropsychologia*, *105*, 70–83. https://doi.org/10.1016/j.neuropsychologia.2017.01.001

Jamrozik, A., & Gentner, D. (2020). Relational labeling unlocks inert knowledge. *Cognition*, *196*, 104146. https://doi.org/10.1016/j.cognition.2019.104146

Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Raichle, M. E., Cohen, N. J., Ollinger, J. M., Akbudak, E., Conturo, T. E., Snyder, A. Z., & Petersen, S. E. (1998). Hemispheric specialization in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron*, *20*(5), 927–936. https://doi.org/10.1016/S0896-6273(00)80474-2

Krawczyk, D. C. (2012). The cognition and neuroscience of relational reasoning. *Brain Research*, *1428*, 13–23. https://doi.org/10.1016/j.brainres.2010.11.080

Krawczyk, D. C., McClelland, M. M., & Donovan, C. M. (2011). A hierarchy for relational reasoning in the prefrontal cortex. *Cortex*, *47*(5), 588–597. https://doi.org/10.1016/j.cortex.2010.04.008

Krawczyk, D. C., McClelland, M. M., Donovan, C. M., Tillman, G. D., & Maguire, M. J. (2010). An fMRI investigation of cognitive stages in reasoning by analogy. *Brain Research*, *1342*, 63–73. https://doi.org/10.1016/j.brainres.2010.04.039

Krawczyk, D. C., Morrison, R. G., Viskontas, I., Holyoak, K. J., Chow, T. W., Mendez, M. F., Miller, B. L., & Knowlton, B. J. (2008). Distraction during relational reasoning: The role of prefrontal cortex in interference control. *Neuropsychologia*, *46*(7), 2020–2032. https://doi.org/10.1016/j.neuropsychologia.2008.02.001

Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, *12*(5), 477–485. https://doi.org/10.1093/cercor/12.5.477

Laird, J. E., Kinkade, K. R., Mohan, S., & Xu, J. Z. (2012). *Cognitive robotics using the soar cognitive architecture*.

Luo, Q., Perry, C., Peng, D., Jin, Z., Xu, D., Ding, G., & Xu, S. (2003). The neural substrate of analogical reasoning: An fMRI study. *Cognitive Brain Research*, *17*(3), 527–534. https://doi.org/10.1016/S0926-6410(03)00167-8

Manenti, R., Cotelli, M., Calabria, M., Maioli, C., & Miniussi, C. (2010). The role of the dorsolateral prefrontal cortex in retrieval from long-term memory depends on strategies: A repetitive transcranial magnetic stimulation study. *Neuroscience*, *166*(2), 501–507. https://doi.org/10.1016/j.neuroscience.2009.12.037

Matsumoto, A., Soshi, T., Fujimaki, N., & Ihara, A. S. (2021). Distinctive responses in anterior temporal lobe and ventrolateral prefrontal cortex during categorization of semantic information. *Scientific Reports*, *11*(1), 13343. https://doi.org/10.1038/s41598-021-92726-7

McLure, M. D., Friedman, S. E., & Forbus, K. D. (2015). *Extending analogical generalization with near-misses.*

Morrison, R. G., Holyoak, K. J., Troung, B., & Truong, B. (2001). *Working-memory modularity in analogical reasoning*.

Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L., & Knowlton, B. J. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience*, *16*(2), 260–271. https://doi.org/10.1162/089892904322984553

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268. https://doi.org/10.3758/s13415-011-0083-5

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430. https://doi.org/10.1016/j.tics.2006.07.005

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*(2), 109–130. https://doi.org/10.1017/S0140525X08003543

Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research*.

Petrides, M., & Pandya, D. N. (2002a). Association pathways of the prefrontal cortex and functional observations. *Principles of Frontal Lobe Function*, *1*, 31–50.

Petrides, M., & Pandya, D. N. (2002b). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291–310.

Power, J. D., & Petersen, S. E. (2013). Control-related systems in the human brain. *Current Opinion in Neurobiology*, *23*(2), 223–228. https://doi.org/10.1016/j.conb.2012.12.009

Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's progressive matrices test. *Cognitive Psychology*, *33*(1), 43–63. https://doi.org/10.1006/cogp.1997.0659

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, *14*(2), 148–152. https://doi.org/10.1002/hipo.20009

Ptak, R. (2012). The frontoparietal attention network of the human brain: Action, saliency, and a priority map of the environment. *Neuroscientist*, *18*(5), 502–515. https://doi.org/10.1177/1073858411409051

Qiu, J., Li, H., Chen, A., & Zhang, Q. (2008). The neural basis of analogical reasoning: An event-related potential study. *Neuropsychologia*, *46*(12), 3006–3013. https://doi.org/10.1016/j.neuropsychologia.2008.06.008

Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*(1), 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Reber, T., Luechinger, R., Boesiger, P., & Henke, K. (2014). Detecting analogies unconsciously. *Frontiers in Behavioral Neuroscience*, *8*, 9. https://doi.org/10.3389/fnbeh.2014.00009

Reber, T. P., Luechinger, R., Boesiger, P., & Henke, K. (2012). Unconscious relational inference recruits the hippocampus. *Journal of Neuroscience*, *32*(18), 6138–6148. https://doi.org/10.1523/JNEUROSCI.5639-11.2012

Renoult, L., Irish, M., Moscovitch, M., & Rugg, M. D. (2019). From knowing to remembering: The semantic–episodic distinction. *Trends in Cognitive Sciences*, *23*(12), 1041–1057. https://doi.org/10.1016/j.tics.2019.09.008

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*(3), 249–273. https://doi.org/10.1016/j.jecp.2006.02.002

Robbins, S. E. (2017). Analogical reminding and the storage of experience: The paradox of Hofstadter–Sander. *Phenomenology and the Cognitive Sciences*, *16*(3), 355–385. https://doi.org/10.1007/s11097-016-9456-0

Rojkova, K., Volle, E., Urbanski, M., Humbert, F., Dell'Acqua, F., & Thiebaut de Schotten, M. (2016). Atlasing the frontal lobe connections and their variability due to age and education: A spherical deconvolution tractography study. *Brain Structure and Function*, *221*(3), 1751–1766. https://doi.org/10.1007/s00429-015-1001-3

Schmidt, G. L., Cardillo, E. R., Kranjec, A., Lehet, M., Widick, P., & Chatterjee, A. (2012). Not all analogies are created equal: Associative and categorical analogy processing following brain damage. *Neuropsychologia*, *50*(7), 1372–1379. https://doi.org/10.1016/j.neuropsychologia.2012.02.022

Smallwood, J., Bernhardt, B. C., Leech, R., Bzdok, D., Jefferies, E., & Margulies, D. S. (2021). The default mode network in cognition: A topographical perspective. *Nature Reviews Neuroscience*, *22*(8), 503–513. https://doi.org/10.1038/s41583-021-00474-4

Stavy, R. (1991). Using analogy to overcome misconceptions about conservation of matter. *Journal of Research in Science Teaching*, *28*(4), 305–313. https://doi.org/10.1002/tea.3660280404

Steinbeis, N., Bernhardt, B. C., & Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron*, *73*(5), 1040–1051. https://doi.org/10.1016/j.neuron.2011.12.027

Thagard, P. (1988). Dimensions of analogy. In D. H. Helman (Ed.), *Analogical reasoning: Perspectives of artificial intelligence, cognitive science, and philosophy* (pp. 105–124). Netherlands: Springer. https://doi.org/10.1007/978-94-015-7811-0_5

Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence*, *46*(3), 259–310. https://doi.org/10.1016/0004-3702(90)90018-U

Thibaut, J. P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development*, *38*, 10–26. https://doi.org/10.1016/j.cogdev.2015.12.002

Urbanski, M., Bréchemier, M.-L., Garcin, B., Bendetowicz, D., Thiebaut de Schotten, M., Foulon, C., Rosso, C., Clarençon, F., Dupont, S., Pradat-Diehl, P., Labeyrie, M.-A., Levy, R., & Volle, E. (2016). Reasoning by analogy requires the left frontal pole: Lesion-deficit mapping and clinical implications. *Brain*, *139*(6), 1783–1799. https://doi.org/10.1093/brain/aww072

Vartanian, O. (2012). Dissociable neural systems for analogy and metaphor: Implications for the neuroscience of creativity. *British Journal of Psychology*, *103*(3), 302–316. https://doi.org/10.1111/j.2044-8295.2011.02073.x

Vendetti, M. S., & Bunge, S. A. (2014). Evolutionary and developmental changes in the lateral frontoparietal network: A little goes a long way for higher-level cognition. *Neuron*, *84*(5), 906–917. https://doi.org/10.1016/j.neuron.2014.09.035

Viskontas, I. V., Morrison, R. G., Holyoak, K. J., Hummel, J. E., & Knowlton, B. J. (2004). Relational integration, inhibition, and analogical reasoning in older adults. *Psychology and Aging*, *19*(4), 581. https://doi.org/10.1037/0882-7974.19.4.581

Vitello, S., Warren, J. E., Devlin, J. T., & Rodd, J. M. (2014). Roles of frontal and temporal regions in reinterpreting semantically ambiguous sentences. *Frontiers in Human Neuroscience*, *8*, 530. https://doi.org/10.3389/fnhum.2014.00530

Volle, E., Gilbert, S. J., Benoit, R. G., & Burgess, P. W. (2010). Specialization of the rostral prefrontal cortex for distinct analogy processes. *Cerebral Cortex*, *20*(11), 2647–2659. https://doi.org/10.1093/cercor/bhq012

Wais, P. E., Kim, O. Y., & Gazzaley, A. (2012). Distractibility during episodic retrieval is exacerbated by perturbation of left ventrolateral prefrontal cortex. *Cerebral Cortex*, *22*(3), 717–724. https://doi.org/10.1093/cercor/bhr160

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., Thomas, C. R., & Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10*(2), 119–125. https://doi.org/10.1111/1467-9280.00118

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory and Cognition*, *28*(7), 1205–1212. https://doi.org/10.3758/BF03211821

Watson, C. E., & Chatterjee, A. (2012). A bilateral frontoparietal network underlies visuospatial analogical reasoning. *Neuroimage*, *59*(3), 2831–2838. https://doi.org/10.1016/j.neuroimage.2011.09.030

Weaverdyck, M. E., Lieberman, M. D., & Parkinson, C. (2020). Tools of the trade multivoxel pattern analysis in fMRI: A practical introduction for social and affective neuroscientists. *Social Cognitive and Affective Neuroscience*, *15*(4), 487–509. https://doi.org/10.1093/scan/nsaa057

Wendelken, C. (2015). Meta-analysis: How does posterior parietal cortex contribute to reasoning? *Frontiers in Human Neuroscience*, *8*, 1–11. https://doi.org/10.3389/fnhum.2014.01042

Wendelken, C., Nakhabenko, D., Donohue, S. E., Carter, C. S., & Bunge, S. A. (2008). "Brain is to thought as stomach is to ??": Investigating the role of rostrolateral prefrontal cortex in relational reasoning. *Journal of Cognitive Neuroscience*, *20*(4), 682–693. https://doi.org/10.1162/jocn.2008.20055

Wendelken, C., Ferrer, E., Whitaker, K. J., & Bunge, S. A. (2016). Fronto-parietal network reconfiguration supports the development of reasoning ability. *Cerebral Cortex*, *26*(5), 2178–2190. https://doi.org/10.1093/cercor/bhv050

Westphal, A. J., Reggente, N., Ito, K. L., & Rissman, J. (2016). Shared and distinct contributions of rostrolateral prefrontal cortex to analogical reasoning and episodic memory retrieval. *Human Brain Mapping*, *37*(3), 896–912. https://doi.org/10.1002/hbm.23074

Wharton, C. M., Grafman, J., Flitman, S. S., Hansen, E. K., Brauner, J., Marks, A., & Honda, M. (2000). Toward neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology*, *40*(3), 173–197. https://doi.org/10.1006/cogp.1999.0726

Whitaker, K. J., Vendetti, M. S., Wendelken, C., & Bunge, S. A. (2018). Neuroscientific insights into the development of analogical reasoning. *Developmental Science*, *21*(2), e12531. https://doi.org/10.1111/desc.12531

Winecoff, A., Clithero, J. A., Carter, R. M., Bergman, S. R., Wang, L., & Huettel, S. A. (2013). Ventromedial prefrontal cortex encodes emotional value. *Journal of Neuroscience*, *33*(27), 11032–11039. https://doi.org/10.1523/JNEUROSCI.4317-12.2013

Wright, S., Matlen, B., Baym, C., Ferrer, E., & Bunge, S. (2008). Neural correlates of fluid reasoning in children and adults. *Frontiers in Human Neuroscience*, *2*, 8. https://doi.org/10.3389/neuro.09.008.2007

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, *75*(1), 168–179. https://doi.org/10.1016/j.neuron.2012.05.010

Zeithamova, D., Mack, M. L., Braunlich, K., Davis, T., Seger, C. A., van Kesteren, M. T. R., & Wutz, A. (2019). Brain mechanisms of concept learning. *Journal of Neuroscience*, *39*(42), 8259–8266. https://doi.org/10.1523/JNEUROSCI.1166-19.2019