CHAPTER

3

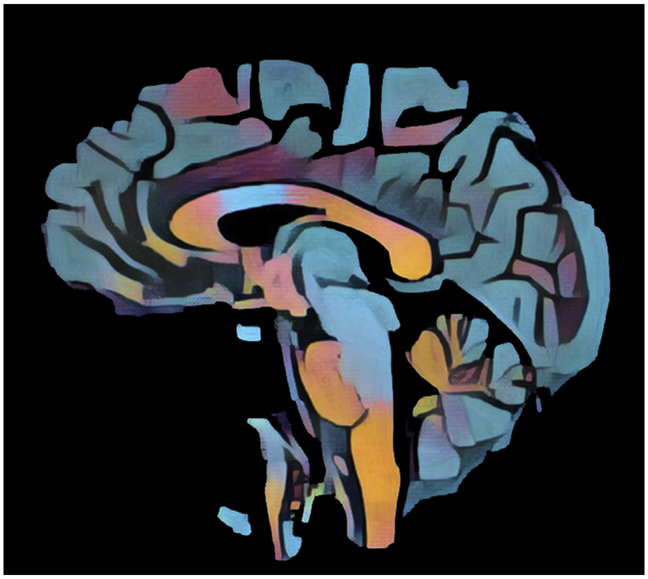
The Neuroscience of Reasoning

KEY THEMES

- Species that possess large amounts of cortex and especially association cortex tend to show advanced reasoning skills.
- The relationship of large brains to complex behaviors may be driven by whether an organism is a predator. Predators need a wide array of behaviors in order to locate and subdue prey.
- Functional neuroimaging involves scanning the brains of participants as they carry out reasoning tasks. This results in the ability to map brain regions that support reasoning.
- Many of the early neuroimaging studies of reasoning suggested that the prefrontal cortex (PFC) was particularly important.
- Relational reasoning and analogical reasoning involve areas of the left PFC. This finding has been indicated by both neuroimaging and electrophysiological studies.
- Neural network studies indicate that the frontal lobes may serve a coordinating or control function. This area may integrate wide-scale activity across the brain, rather than operating as a relational module.
- The materials used in an experiment have a large impact on the results of reasoning studies. Materials invoke our semantic memories, which are supported by temporal lobe regions.
- The integrated sets of information that we can use in reasoning are called schemas or scripts.
- Brain network interconnectivity looks to be an especially promising area toward capturing the complexity of neural processing in reasoning and may further clarify some of the roles of specific brain areas that have been linked to reasoning in various forms.

Reasoning
http://dx.doi.org/10.1016/B978-0-12-809285-9.00003-X

OUTLINE

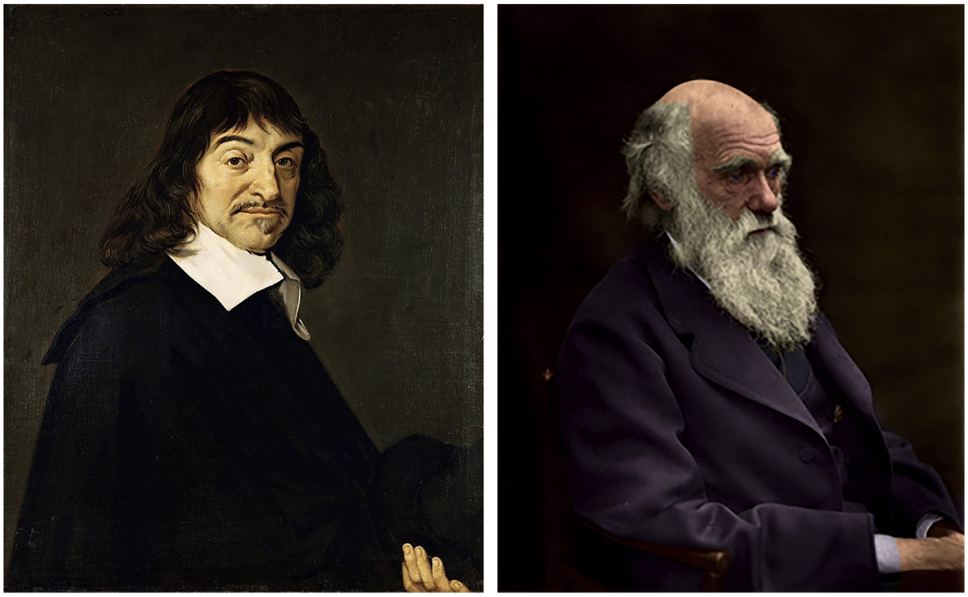


THE NEUROSCIENCE OF REASONING

Introduction

The quest to understand the neural basis of reasoning is a relatively new and complicated endeavor. Researchers working within this subfield of the discipline wish to determine how the cognitive operations of reasoning are carried out within the brain. Achieving this goal would provide some much needed clarity about what categories of reasoning are sufficiently different that they have clear biological differences and rely upon different brain systems. More importantly, a better understanding of how the brain carries out reasoning could guide medical practitioners toward targeting key processes that become impaired in brain injury and dementia patients. A greater understanding of the brain mechanisms for reasoning could also benefit educators as they strive to provide the most fertile learning environments possible at the most developmentally appropriate time.

The history of philosophical inquiry into the nature of reasoning has frequently touched upon this very topic. In the 1600s Rene Descartes famously grappled with mind-body questions. Descartes maintained that reason was separate from the body and that it was the perfect product of the mind. His premises led to the conclusion that the mind is not constrained by biology and that we do not benefit from biological investigations of the mind. By the late 1800s Charles Darwin articulated a different position suggesting that human reasoning and mental abilities exist on a continuum with the abilities of other species and that biological traits have very much to do with reasoning abilities (Fig. 3.1). While this was a highly controversial position in the late 1800s, science would gradually begin to widely adopt this view, which paved the way for educated societies to view other species in a more ethically humane way. Through the 1900s many academic disciplines expanded dramatically including the fields of biology, neuroscience, psychology, neurology, psychiatry, physiology, and anthropology. This expansion has enabled us to begin to integrate across many of these disciplines by gathering evidence



RENE DESCARTES CHARLES DARWIN

FIGURE 3.1 French philosopher Rene Descartes (1596–1650) and British naturalist Charles Darwin (1809–1882) articulated opposing positions separated by several hundred years regarding the nature of the relationship between the mind and body. *From Wikimedia Commons.*

from a variety of techniques from several fields of study. This integrative process is helping us to define the ability to reason and develop plans to facilitate reasoning skills.

One of the most practical applications for a neuroscience-based understanding of reasoning is applying that knowledge toward understanding disorders of thought and reason. We hope to provide better behavioral and biological markers with the potential to help diagnose and treat brain injuries and neurological conditions. This is an ambitious goal, but one that the modern cross-disciplinary emphasis may one day make possible. My own career has been very much shaped by the convergence of numerous subdisciplines on the topic of neuroscience and cognition. I was trained in a psychology department studying cognition and neuroscience. I then completed my training at both a psychology department and an academic neuroscience center. I am currently a faculty member in a school of behavioral and brain sciences and a psychiatry department. I am also affiliated with a university neuroscience center, as well as a medical brain imaging research center. This lack of organizational boundaries can at times be complicated, but ultimately offers excellent opportunities to study reasoning using a variety of tools and methods in the contexts of experimental psychology and neuroscience. The influence of numerous brain-related disciplines interacting may enable us to help patients to regain reasoning abilities through targeted medical and rehabilitative interventions.

A good place to begin our inquiry into the neuroscience of reasoning is considering the differences between people and animals in reasoning ability. While we will focus more deeply on this topic in Chapter 4, it is important to make some essential points before we begin. When we consider humans in the context of other species the definition of reasoning begins to take shape and possibly change. This comparison forces us to clarify what type of behaviors should be considered *reasoning* and opens up a variety of interesting questions. Is it reasoning when birds flock together and migrate? Is it reasoning when an ant colony works as a group to find food and build a home? Is it reasoning when a trout deftly eats your bait while avoiding getting caught on your fishing hook? Are these instead instances of *learning*? Might some of these behaviors be driven by *instinct* due to the strong genetic guidance? Alternatively, must an organism be creative in a behavior in order for it to count as reasoning, or can reasoning be merely an elaborate form of imitation drawing from the abilities of one's family members? To answer these types of questions, we must consider how the behavior is represented at the level of the brain. This leads us to a separate, but related issue regarding the role of brain size, capacity, and the characteristics of an organism's nervous system. These critical features are intertwined with the reasoning abilities of an organism.

Differences in Brain Capacity

Both the size and characteristics of a nervous system are important determinants of reasoning abilities. As a general principle, larger brains usually lead to more complex behavior, and some of this complex behavior might be defined as reasoning. A challenge we quickly run up against is arriving at a precise definition of reasoning in other species. Reasoning is clearly a human construct that we have applied to several of our own abilities, but how should we define the abilities of animals? We will revisit this topic in the next chapter when we cover comparative reasoning abilities in depth, but for now it may be sufficient to acknowledge some broad trends about the links between reasoning and the brain.

Overall brain size is an important indicator of both brain complexity and behavioral complexity. Many of the largest animals on Earth have large brains. Many of these are mammals that have relatively complex brains with pronounced anatomical and physiological similarities that can be observed across species. We can also observe this trend based on evidence from extinct species. Several species of dinosaurs were enormous by modern day standards, but their heads were small relative to the size of their bodies. Estimates of brain size in some of these species indicate that they had very small brains compared to their body size. The 30 ft long 3.5-ton stegosaurus famously had a brain about the size of a walnut, while the much smaller Troodon had a relatively large brain with a larger proportion of cerebrum, the area comprised of the cortex (Fig. 3.2). The behavioral traits of these animals can only be inferred from the fossil record, but paleontologists suggest that the Troodon was a fast hunter with stereoscopic vision, while the stegosaurus was likely a slower moving grazing animal. Large modern animals can be compared in terms of brain size as well, with whales, dolphins, and elephants emerging as some of the most intelligent, large-bodied, large-brained species. All of these features tend to result in more complex behaviors.

Another strong predictor of reasoning ability and brain size is whether an animal is a herbivore or predator. Herbivores include many large animals such as giraffes, buffalo, bison, moose, elk, and zebras. Despite the large body size of these animals, none of them have especially large brains. Meanwhile, carnivores such as cheetahs, bobcats, and hyenas are much smaller animals that have proportionally large brains. Not only do the predators tend to have larger brains, they are often capable of a richer set of diverse behaviors compared to grazing herbivores. Perhaps these differences can be attributed to the ways these animals obtain food. The daily activities of herbivores include eating, migrating, sleeping, and monitoring for the possibility of danger. If no danger presents itself, they have a relatively peaceful day. By contrast, predators must identify, capture, and subdue prey animals by overcoming their defenses. Prey animals may also be spread over vast territory and may be hidden. These additional pressures likely drive the evolution of enhanced brain capacity in predators.

The proportion of cortical mass to brain size is strongly linked to reasoning abilities. Humans have a remarkably large proportion of cortex relative to body size. Similarly, large cortical mass is a prominent feature of cats, dogs, and many other predatory species. Predators tend to have binocular vision as well, which enables them to better estimate depth cues in the environment. Such abilities are made possible by complex nervous systems that have the processing capacity necessary to model an accurate representation of the visual landscape. Similarly, dolphins and killer whales use sound to augment their sensory representations of the oceanic world in which they live. Predators that must track and pursue single individual prey animals often require sophisticated sensory abilities, and these sensory capacities tend to accompany complex brains. It is also important to note that many species of birds are capable of behaviors that many would classify as involving reasoning skills. Such behaviors include the use of tools, a level of sensitivity to the context, and the ability to process abstract relations among objects. Birds have physically small brains compared to large mammals, but their brain size is proportionally large relative to their small bodies. The proportional size

FIGURE 3.2 We can infer the complexity of behavior in extinct species by considering clues in the fossil record. The troodon was a dinosaur and likely a clever hunter with a relatively large brain. By contrast the stegosaur was a relatively large animal that had a very small brain. *By Reid, Iain James (Drawn by hand) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0), via Wikimedia Commons.*

between the brain and the body is often a better indicator of how sophisticated the reasoning abilities and complex behaviors of an organism are.

ANATOMICAL CONSIDERATIONS

The Importance of Association Cortex for Higher Cognitive Functions

The cortex, or outer surface of the brain, contains the cell bodies of neurons, which are the primary computational units within the brain. The overall amount of cortex is often related to the complexity of behavior that we observe in a given species. More processing power made possible by greater amounts of cortical tissue could be an important feature of the brain that relates to the complexity of reasoning abilities; however, there are likely some other key biological factors that bear on an individual's ability to reason.

Cortical areas rarely act as isolated modules or units. Instead regions of the cortex tend to communicate in networks. Primate memory physiologist Joaquin Fuster has referred to these cortical networks as *cognits* (Fuster, 2006). The idea that cortical areas are critically interconnected becomes apparent when one considers the structure and functions of single neurons. Most neurons will have numerous dendrites, which are the projections that enable them to receive chemical or electrical information from other neurons. Neurons also typically have axons, which enable the flow of electrochemical reactivity across space within the brain. Lastly, the terminal buttons projecting off of the axon allow neurons to communicate with one another. Thus, it should not be surprising that the types of connections possessed by neurons or cortical areas will lead to large-scale network interactions and to interactions of more localized circuits (Fig. 3.3). It is likely that those network interactions that can span across the hemispheres, from front-to-back, or up-to-down, are especially important for integrating information needed for complex reasoning skills.

In addition to the mass and connectivity of the cortex, the structure of the brain is also important to consider. Some cortical areas such as the primary visual cortex, located in the occipital lobe, receive relatively unprocessed information from the retinas. The neurons within this area of the brain are tuned for detecting edges, features, and differences in lighting. A similar arrangement is found within Heschl's gyrus on the top of the temporal lobes, which is the primary cortical area associated with processing sound. The neurons here, like those in the visual cortex, are finely tuned to respond to different incoming sound frequencies transduced into electrochemical information by the inner ear. The primary motor cortex, which lies just forward of the central sulcus, contains neurons that directly link to the outgoing axons that extend out to the periphery of the body. These neurons are predominantly tuned for sending out motor signals. These primary sensory and motor areas are not limited only to these functions, but a large percentage of the neurons in those areas are heavily involved in sensory and motor functions. These areas may participate in activities that we would call reasoning, but they probably play limited roles bound to information either entering or exiting the brain having to do with the format or modality of the reasoning problem.

Some of the important neurons that are involved in reasoning are those that comprise the *heteromodal cortex*. This term is used to describe the cortex that does not serve one particular function, but rather participates in

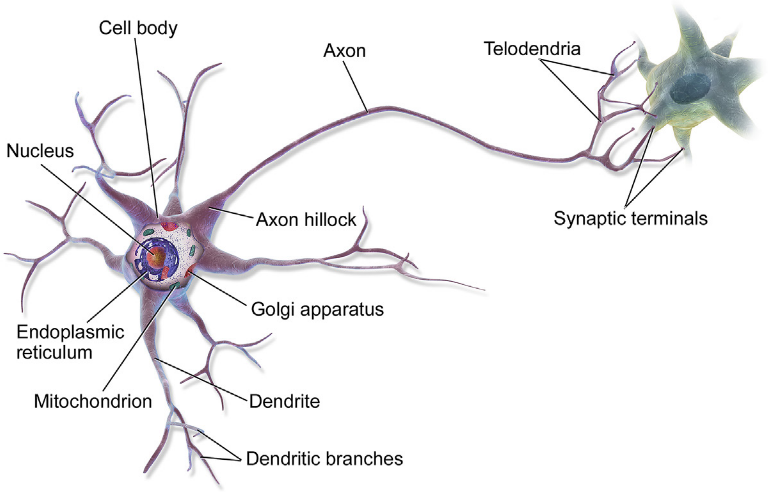
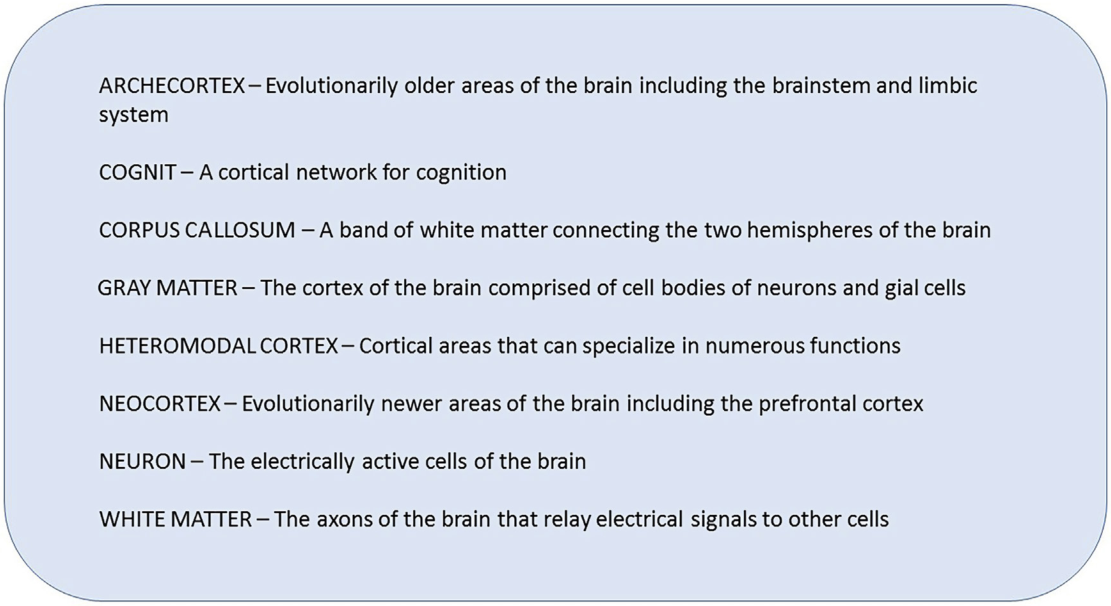


FIGURE 3.3 The structure of a neuron includes the axon, dendrites, cell body, and synaptic terminals. The dendrites, axons and terminal buttons enable neurons to connect to form large-scale and small-scale networks. *CC BY-SA 3.0, from Wikimedia Commons.*



ARCHECORTEX	Evolutionarily older areas of the brain including the brainstem and limbic system
COGNIT	A cortical network for cognition
CORPUS CALLOSUM	A band of white matter connecting the two hemispheres of the brain
GRAY MATTER	The cortex of the brain comprised of cell bodies of neurons and gial cells
HETEROMODAL CORTEX	Cortical areas that can specialize in numerous functions
NEOCORTEX	Evolutionarily newer areas of the brain including the prefrontal cortex
NEURON	The electrically active cells of the brain
WHITE MATTER	The axons of the brain that relay electrical signals to other cells

FIGURE 3.4 Neuroanatomical terms help us to describe aspects of the nervous system.

many functions. Such areas are also called association cortex, as they tend to both receive information from many brain areas and project information to many other areas. Association cortex, or heteromodal cortex, may be in an ideal position to participate in guiding our reasoning functions, as these areas can both receive and relay larger integrative aspects of information processing (Fig. 3.4).

The Expansion of the Association Cortex in Predators

Brains and nervous systems have expanded as organisms have evolved new behaviors that are better tuned to enhance their survival skills. Some of the most widely shared anatomical areas of the brain across species are referred to as *archecortex*, meaning “old cortex.” Areas of the archecortex include the limbic system structures, such as the hippocampus and amygdala, the brainstem nuclei, thalamus, and hypothalamus related to regulation of body states. Archecortical areas tend to be involved in primary survival functions. These include regulating the heart rate and respiratory system, coordinating sensory information, and allowing for emotional reactivity toward environmental variables. Archecortical areas tend to be similar across a variety of species. For example, small mammals such as mice have remarkably similar hippocampi, amygdalae, and thalami compared to primates and larger mammals. All of these areas appear to have similar functional roles across a variety of mammals, and they are also located at centralized positions often beneath the later-evolved areas of the cortex.

Association cortex is often referred to as the *neocortex* or “new cortex” due to its relatively late appearance in the evolutionary history of organisms. The neocortex comprises mostly of the outer covering of the brain in large and complex mammals including primates, elephants, dolphins, and whales. Unlike the archecortex, the neocortex tends to lack clearly defined single functions, or even well-defined sets of functions. Rather, neocortical areas appear to become involved in many diverse functions that are carried out by an organism. Such functions may be so difficult to specify that we simply refer to the supporting brain areas as “integrators” that receive highly processed information from other areas of the nervous system. These integrator regions perform processing functions, and then route their outputs to other areas of the nervous system that are downstream and closer to the ultimate action or output of behavior by the organism. The neocortex includes most of the frontal lobes in primates, including the prefrontal cortex (PFC), which is dramatically expanded within the primates. Anatomists currently consider humans and chimpanzees to possess the largest PFC proportionally to body size within the mammals. The PFC is highly important for reasoning functions as we will see in this chapter.

Another important neocortical area is the parietal cortex. The parietal lobes are densely interconnected with the PFC, as well as with the temporal lobes, and occipital cortex. Functions of the parietal cortex include a role in calculating numerical quantities, locating objects in space, coordinating grasp for grasping objects, orienting our hands, and visualizing spatial layouts. The PFC and parietal cortex also show a high degree of correlation. Indeed the frontal-parietal network has been identified as one of the most commonly active task-based networks within the brain through studies of functional brain connectivity. As we will describe in this chapter, the parietal cortex is frequently co-active with the PFC in carrying out reasoning tasks.

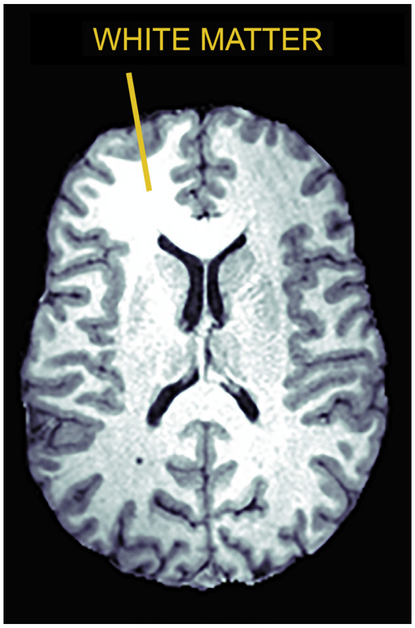


FIGURE 3.5 The neurons of the brain are densely interconnected through bundles of axons that are collectively known as tracts or fasciculi. These are commonly known as the white matter of the brain and this tissue images white in MRI scans.

Communication Across Cortical Areas

Cortical association areas are densely interconnected with the rest of the brain. Many of these connections form the large crossing fiber bands of the corpus callosum, the wide-ranging longitudinal fasciculi connecting frontal, occipital, and parietal cortex, and the U-fiber bands which allow multiple cortical regions to communicate (Fig. 3.5). Even in cases where there is not a single tract or bundle of axons connecting the cortical areas, communication can still be achieved through synapses at the thalamus, which is comprised of a set of nuclei within the midbrain that has a characteristic density of incoming and outgoing fiber tracts reaching across vast cortical territory. The thalamus is sometimes considered to be a relay station capable of routing messages from distant cortical areas to other areas through synapses within its subnuclei. My esteemed neuroanatomy instructor Arnold B. Scheibel would refer to the thalamus as the "gateway to the cortex."

NEUROSCIENCE OF RELATIONAL REASONING

From Hypotheses to Neural Network Models and Experiments

In the 1980s researchers became increasingly interested in modeling cognitive operations in ways that make use of the principles of neural systems. For example, the connectionist parallel-distributed processing (PDP) models pioneered by David Rumelhart and James McClelland (1986) were capable of modeling a range of complex cognitive abilities including perception, reading, and reasoning. These models were termed "biologically plausible," as they featured a series of primitive processing units in the form of positively and negatively weighted units. These units that would share connections with other layered units and simulated energy (representing neurochemically based signaling) were transferred through the connections to produce more sophisticated output from the models.

By the 1990s researchers had become increasingly interested in how actual nervous systems functioned and fewer papers appeared reporting comparisons between human behavior and the connectionist architectures. Some researchers began to describe the approach as being "neurally inspired" rather than "neurally plausible," as no connectionist model was able to lay claim to realizing the goal of a faithful reproduction of an actual nervous system. All such models were extremely oversimplified at the level of connectivity. This led them to be taken less seriously by the research community that was building what are known as "neuronal models." A neuronal model is not aimed at simulating cognitive phenomena, but rather it is meant to model simple neural circuits with high fidelity. Such models might be used to reflect the cell interactions that make up a reflex response in an animal with a parsimonious nervous system. These much simpler neural circuit models could more accurately be characterized as being "neurally plausible."

In the middle of the 1990s several new types of neuroimaging tools were becoming available. Among these were positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) (refer to Box 3.1). These techniques are capable of imaging brain tissue as a person performs a sensory, motor, or cognitive task while having their brain scanned. PET and fMRI allowed researchers to plot functional activity onto anatomical scans, thus mapping out and isolating key areas that were involved in different types of tasks at the scale of the entire brain.

In 1995, Nina Robin and Keith Holyoak published a book chapter in which they hypothesized that the PFC was likely to be an important area for carrying out the operations of human reasoning. Robin and Holyoak described lines of evidence from brain organization, including the fact that the frontal lobes were highly interconnected with many other brain areas, that patients with frontal damage would exhibit challenges in maintaining items in working memory, avoiding distractions, and had difficulties with focusing attention. This was one of the first reports in the reasoning literature articulating the possibility that the frontal lobes may be the most critical region of the brain for relational reasoning or integrating information across dimensions. With this hypothesis, the new functional brain imaging methods could be deployed alongside more traditional investigations of cell-recordings in nonhuman primates, as well as patient studies of individuals with brain damage to test whether the PFC was central to relational reasoning.

Investigating Reasoning Ability in the Brain

One of the experimental tools of choice for evaluations of relational reasoning is the Raven's matrices task (Raven, 1938, 1960). In the task, a person sees a series of abstract visual objects arranged in a three-by-three matrix with one item missing. The goal of the participant is to determine what object or visual pattern best fits into the missing space by evaluating the patterns contained within the other places in the matrix (Fig. 3.6). The simplest matrices problems involve basic perceptual matching, in which the item to complete the problem matches based on appearance alone. More complex Raven's matrix problems require the participant to evaluate the different types of changes that occur either across the rows or down the columns and select the object that best matches the pattern of changes or combinations of

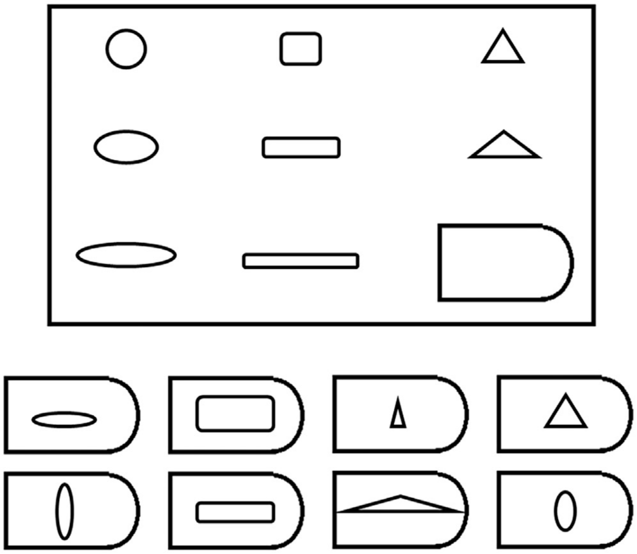


FIGURE 3.6 The Raven's progressive matrices task presents a series of abstract geometric shapes. The goal of the participant is to find the missing piece that best completes the pattern present in the matrix from a set of answer choices.

BOX 3.1

FUNCTIONAL MAGNETIC RESONANCE IMAGING

Magnetic resonance imaging (MRI) is a technique that allows researchers to view internal tissue using the principles of nuclear magnetic resonance. MRI scanners are large devices containing magnetic gradient devices that can generate a magnetic field many times that of the Earth's magnetic field. Magnetic energy can be deposited into the human body using a send and receive coil, which is a sophisticated device enabling both the deposition of energy into a person and the reception of information about the tissue. Researchers can create a picture of any tissue within the body provided that the tissue has composition differences. For example, bone and muscle image vary differently on an anatomical MRI, which is a static picture showing human anatomical structures. Functional MRI uses the same device and the same principles to image the differences between brain activity recorded under experimental conditions.

MRI was first used to image a human in July of 1977 in the physics lab of Raymond Damadian. Early scanners were used to image anatomy at relatively crude levels. The first MRI image was taken of the human chest. It took researchers approximately 5h to generate the first image, which was crude by modern standards. MRI technology improved to the point of becoming a dominant tool for the diagnosis of both disease and injury by the 1980s. In 1990, Seiji Ogawa first demonstrated that MRI could be used to track the activity associated with brain function, not simply structure.

Tracking functional activity was made possible by several key properties of nuclear magnetic resonance and brain physiology. First, neurons produce action potentials at a baseline rate when a particular set of neurons are operating at a resting level. When the brain responds to cognitive or motor demands, the associated neurons tend to activate above their baseline levels producing more action potentials per second. Oxygen in the immediate surrounding areas begins to be depleted in the wake of these more active neurons leading to the presence of more deoxyhemoglobin (deoxygenated blood) relative to the baseline state. In response to this decrease in oxygen, the cardiovascular system delivers an even larger supply of oxyhemoglobin, or oxygenated blood, to the area around the activated neurons. This change in the proportions of local oxyhemoglobin and deoxyhemoglobin leads to a signal that can then be tracked using MRI. The specific MR signal that occurs in this case is called the blood oxygen level dependent (BOLD) signal, and it has been the primary measure used in most of the fMRI studies conducted over the past 20 years. Functional MRI data consist of a series of voxels, which are three-dimensional pixels. There are thousands of voxels acquired across the whole brain during an fMRI experiment, usually obtaining a new reading of the activation (through BOLD difference) every 1 to 2 s.

BOX 3.1 (*cont'd*)

When people design fMRI tasks, they typically include one or more control conditions in which some aspect of interest is carefully controlled for. An example of this occurs with fMRI studies of working memory in which a set of letters to be kept track of is presented to a person in the MRI scanner. In addition to a letter memory condition, researchers will often include a control condition in which letters are shown with a delay along with a test item, but without the memory demand. This enables the researcher to perform a statistical comparison between the memory condition and the control (non-memory) condition. Whichever voxels are more active during the memory condition over the control condition are considered to be relevant to the operations of working memory.

Functional MRI scans often consist of activation maps. These maps are color coded to reflect the degree of average BOLD signal contrast between conditions within the voxels of the brain. Typically, densely clustered groups of neurons will become active together and this will lead to a set of grouped clusters of voxels that are considered to be active over their baseline state. The maps of lower resolution voxels (typically 3 mm cubed) can then be plotted onto a higher resolution structural MRI image of the brain allowing the researcher to view the areas of activity accompanying a task (Fig. 3.7).

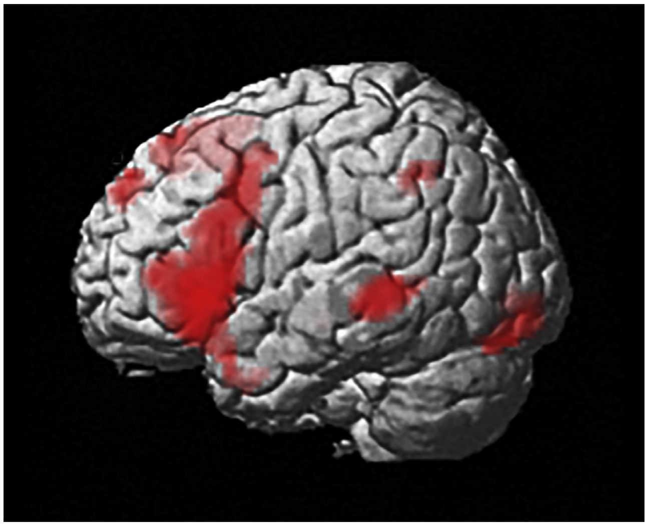


FIGURE 3.7 A color coded map showing fMRI data plotted in red on an anatomical brain scan shown in gray.

changes across the other items. This is a classic relational reasoning task, as it demands evaluation of a series of relationships among the items. The Raven's matrix task is considered to be a culture-fair test, as it lacks any specific identifiable objects. Each figure is a novel geometric shape or pattern. This means that the individualized knowledge, or semantic memory of the individual, is unlikely to affect performance. This task has also been considered to be a test of fluid general intelligence.

Localization of Reasoning and the Prefrontal Cortex

Prabhakaran, Smith, Desmond, Glover, and Gabrieli (1997) at Stanford University performed an early neuroimaging study of reasoning. This study was the first of a line of investigations, in which human participants were scanned using either PET or fMRI while they performed a variation of the Raven's progressive matrices task. In this case, Prabhakaran et al. offered participants figural problems, in which participants evaluated a single dimension of change across a matrix in order to choose a match to complete the pattern from among a set of alternatives. They also offered participants analytic matrix problems, in which multiple relations among task elements had to be integrated in order to solve the problems correctly. This analytic condition provided the best neuroimaging test of the hypothesis that Robin and Holyoak had articulated 2 years earlier, which had stated that the PFC would likely be key to the integration of relational information. The study revealed that bilateral frontal areas within the dorsolateral prefrontal cortex (DLPFC) were more active for the analytic problems compared to the figural problems. Additional areas showing greater activation for the analytic problems included the left parietal cortex, temporal lobes, and occipital regions. All of these active areas were consistent with the maintenance and manipulation of working memory, in addition to any cognitive operation that could plausibly be described as reasoning. A question remained as to whether any of these areas were specifically engaged by relational reasoning, or whether they were alternatively markers of working memory and attention processes that increased most in the demanding analytic condition.

Soon additional matrix reasoning studies appeared in the growing functional neuroimaging literature. These studies were designed to further clarify the role of the relational processing regions that had been identified in the Prabhakaran et al. (1997) study. Christoff et al. (2001) further examined the role of the PFC in matrix reasoning. In this study, participants solved Raven's matrices problems that varied in the number of relations that had to be integrated in order to solve the problem. Participants in this study chose one item that best completed a nine-cell matrix. Similar to the results reported by Prabhakaran et al. (1997), the study by Christoff et al. revealed activation in the left DLPFC, left rostrolateral prefrontal cortex (RLPFC), and bilateral inferior frontal gyrus when two-relation problem solving was compared to one-relation problems. One challenge in interpreting the results in this study was the fact that two-relation problems were more difficult and therefore took longer for participants to solve. This could mean that the PFC was not responding to relational reasoning load, but rather the amount of effort or processing time that was needed to accomplish the task. Christoff et al. addressed this challenge by performing a response-time corrected model for their fMRI data. This procedure was better able to equate the conditions in terms of the fMRI response analysis. This more stringent timing-corrected analysis revealed two left-sided areas of activation within the PFC. The left posterior PFC was one of these areas, which extended over the inferior frontal gyrus. The second area was located within the left RLPFC. These areas specifically appeared to be involved in the integration of relations among items.

Christoff et al. (2001) then examined the time-course of activation of both right- and left-sided DLPFC and RLPFC regions. They found that only the left RLPFC showed a clearly differentiated response to the two-relation problems over the one- and zero-relation problems (refer to Fig. 3.8). The other PFC areas were active across all relational conditions. The left RLPFC region appeared to be the most engaged area in response to integrating relational information, a key skill needed to solve the more difficult Raven's matrices problems.

At this point it remained unclear whether the left RLPFC functions as a relational processing module or whether it was responding to some other measure of task complexity. Another possibility is that the left RLPFC was supporting the maintenance of relational information in working memory. To further explore the involvement of the RLPFC in relational reasoning, Kroger et al. (2002) conducted an fMRI study, in which participants solved problems ranging from zero relations up to four relations. The task was more relationally demanding than the prior fMRI studies of relational reasoning. Kroger et al. found similar results to the previous studies with relational visuo-spatial problem solving evoking bilateral PFC activation and bilateral parietal lobe activation. The left PFC was more engaged than the right PFC in this study. Progressively more anterior

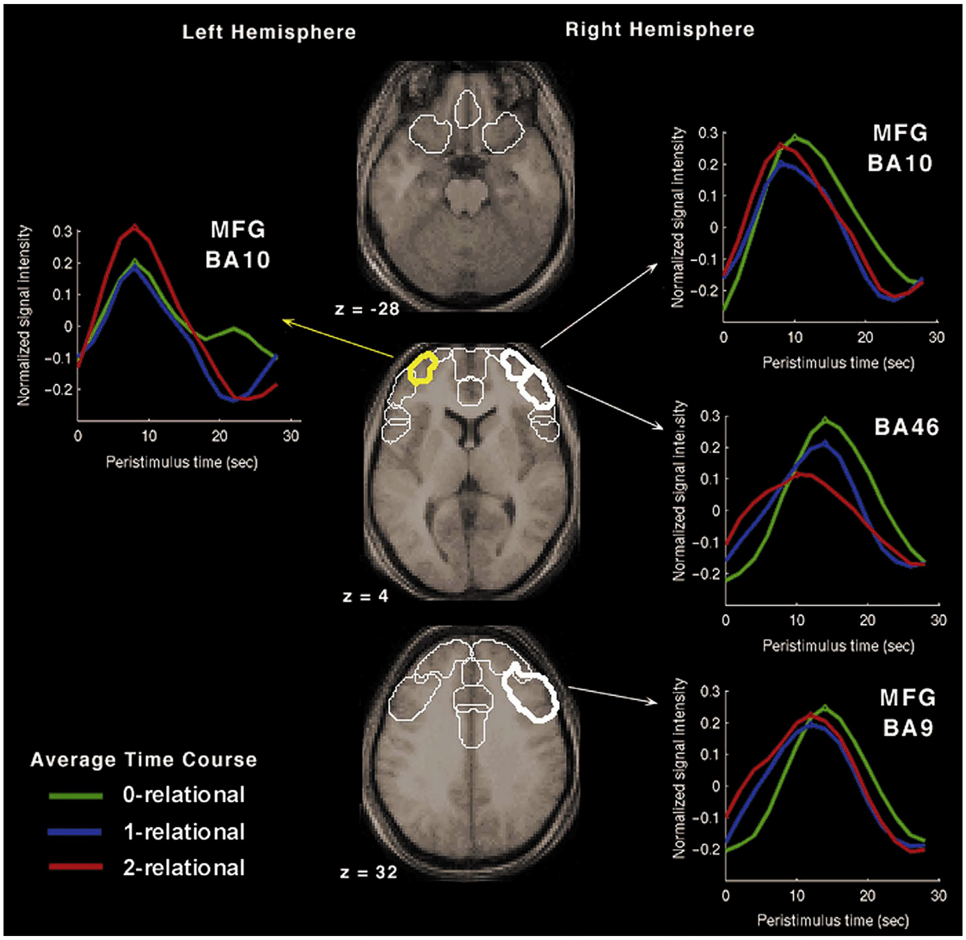


FIGURE 3.8 Christoff et al. (2001) found left PFC activation associated with solving complex relational matrix reasoning problems. *Credit: Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., et al. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning, NeuroImage, 14, 1136–1149.*

regions of the left PFC became activated at higher relational load levels supporting the role of the left RLPFC in relational reasoning.

A serious experimental issue still remained in determining whether the PFC was responsive to relational processing in matrix reasoning tasks. Most neuroimaging studies require a series of simpler control conditions, which will also evoke brain activation patterns of their own. For example, Kroger et al. (2002) included a three-relation condition that was in effect, a control condition for their four-relation condition. This is possible because the activity evoked in the simpler three-relation condition could be removed by subtracting it out from the commonly active areas present at the four-relation level. A key control condition was included in the study by Kroger et al., and this was represented by matrix reasoning problems that were made more difficult and slower to process, but not by the addition of relational elements. Rather, the researchers added visual noise in the form of distractor items that were not relevant to the task. These extra "noise" items were added to all of the relational elements in the matrix. The noise items did not need to be factored into the solution of the problem, but were nonetheless difficult to ignore. These control conditions added processing time and difficulty to the task without increasing the relational demand on the participant. Participants took longer to solve the problems in the conditions with distracting elements added, and this also led to increases in activation in similar bilateral PFC areas, as well as medial PFC regions. Kroger et al. constructed a brain map in which the activation evoked by relational reasoning was compared with that evoked by problem difficulty that was attributable to distracting elements. There was some overlap among areas within the left and right PFC, but only the relational increase led to left RLPFC activity. Furthermore, larger areas of PFC were activated in response to the increased relational load over the distraction factor. Consensus was beginning to build that relational reasoning demanded left RLPFC activity relative to other forms of executive function demand.

There was also interest in Raven's matrices from the intelligence testing community. This group of scholars considered the Raven's matrices task to be part of the construct *fluid intelligence*. Fluid intelligence was a term coined by Raymond Cattell, who had been influenced by the work of Charles Spearman before him. Fluid intelligence refers to the abilities of an individual that are based on biological factors including the sensory structures, the central nervous system, and heredity (Horn & Cattell, 1966). John Duncan et al. (2000) used PET scanning to evaluate the neural systems involved in fluid general intelligence. They boldly titled their paper *A Neural Basis for General Intelligence* and it appeared in the high-impact journal *Science*. Duncan et al. PET scanned participants while they completed three different tasks that contained a range of difficulty (or loading) on fluid general intelligence. For example, participants solved an abstract figure test that had been developed by Raymond Cattell, along with a complex letter sequence reasoning task and another abstract figure judgment task. The high demand conditions within all three of the tasks were accompanied by higher levels of activation within the left DLPFC within the middle frontal gyrus. Duncan et al. argued that the left DLPFC was likely to be a critical region for supporting the mental processes that are necessary for fluid intelligence (Fig. 3.9).

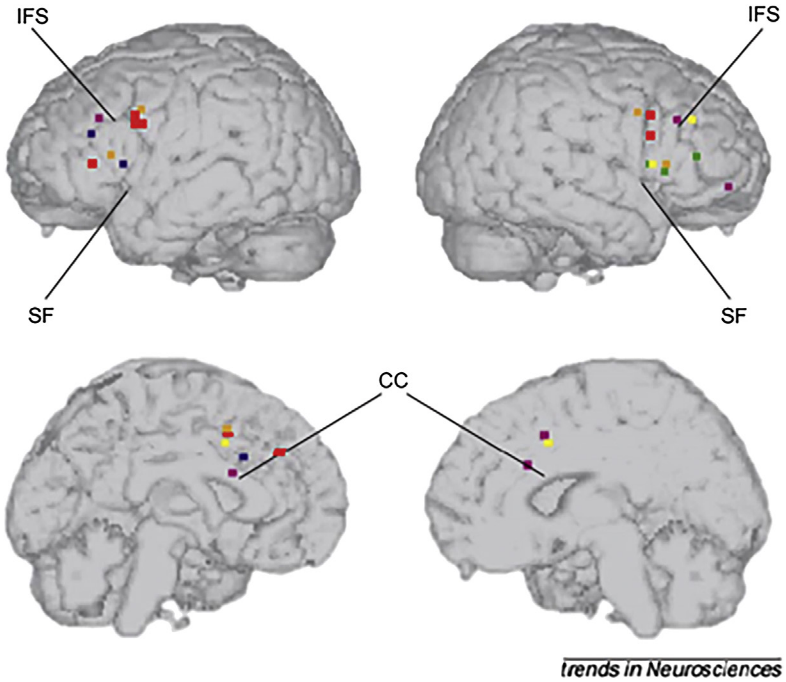


FIGURE 3.9 An influential positron emission tomography (PET) study by Duncan et al. (2000) revealed that a common left frontal region became active across a set of varying fluid reasoning tasks. *Credit*: *Duncan, J., & Owen, A. M. (October 2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends in Neuroscience, 23(10), 475–483.*

The fluid intelligence paper by Duncan et al. (2000) generated quite a bit of enthusiasm, excitement, and discussion, as this was one of the first studies to provide some clues about the biological basis for the construct of fluid intelligence across multiple task formats within a specific region of the brain. The study also generated a fair amount of controversy as well. Noted intelligence theorist and researcher Robert Sternberg, who is well-known for his triarchic theory that described multiple intelligences, wrote a scathing review of the study referring to it as representing the "*holey* grail of intelligence" (Sternberg, 2000). Sternberg expressed his disappointment in the interpretation of the study recommending that the neuroscience community focus on characterization of the cognitive processes that are supported by the PFC, rather than claiming that a concept as difficult to define and variable as intelligence could be supported by a single region of the brain. This critique noted the similarity of this approach to phrenology, a widely criticized effort to map human abilities to particular areas of the head, which was popular in the early 1800s (Fig. 3.10).

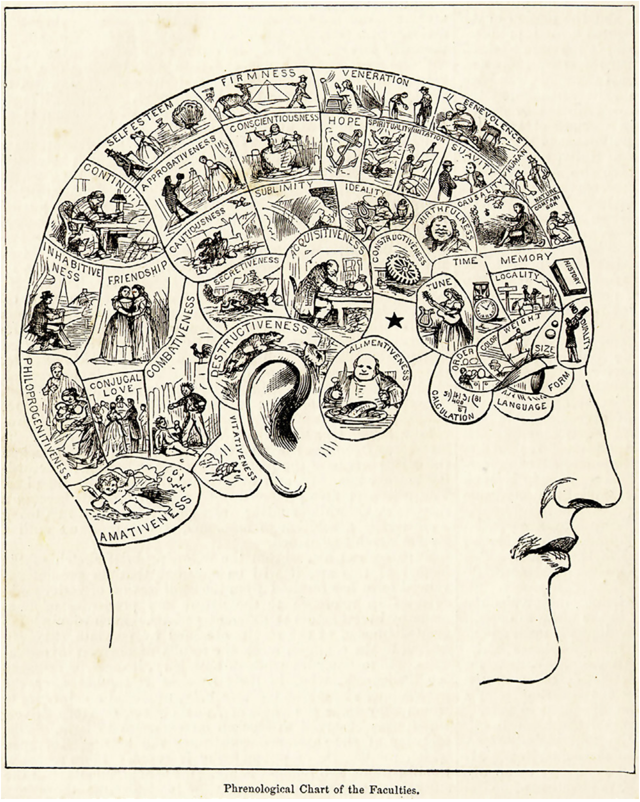


FIGURE 3.10 Phrenology was a movement attempting to link personality and mental capabilities to different regions of the brain. This movement occurred in the 1800s and involved feeling the sizes of areas of the head and concluding that certain propensities were linked to variations in head size. *From Wikimedia Commons.*

Left Prefrontal Cortex and Relational Reasoning With Analogies

The left PFC emerged as important for analogical reasoning very early in the developing neuroimaging literature. Charles Wharton et al. at the National Institutes of Mental Health published one of the first analogical reasoning studies. This was a PET study, in which the analogies consisted of simple shapes that also varied in color and texture. The task was administered in a match-to-sample format. Participants initially encoded a sample set of four shapes followed by a brief delay period. They then had to decide if a second set of four shapes matched the initial sample. In a literal condition, participants simply had to judge from memory whether the two sets of shapes matched. In an analogy condition participants had to judge whether the relations among the four shapes on the test screen matched the relational information present in the sample (refer to Fig. 3.11). Critically, in the analogical match condition the shapes and colors involved in the sample were not the same as those in the match. Rather than match on perceptual features, participants had to understand the similarity between match and sample at a more abstract level. The analogy condition preferentially activated the left parietal cortex, medial frontal cortex, along with a large area within the left PFC. This study indicated a focus of relational processing within the left PFC in an analogical matching condition. The study is reminiscent of the Raven's matrices match studies described earlier in the chapter with some variation in format. Like the matrix reasoning tasks, this analogy task lacked semantic information limiting its ecological validity and generalizability.

A challenge with functional imaging methods are that any inferences one makes are correlational rather than causal. This means researchers cannot infer strong cause-and-effect relationships between cognitive operations and task-evoked brain activation on the basis of those imaging studies alone. There is always the possibility that activity in a given region, such as the left PFC, could be related to some aspect of cognition that accompanies the reasoning construct under examination. For instance, the left PFC may be especially active when people have to maintain many dimensions in working memory within a matrix task. It is also possible that the left PFC is active in reasoning conditions simply because the demand on attention increases in response to the need to notice and consider multiple relational elements. While this second explanation of left PFC involvement would be of great interest to researchers involved in studying reasoning, it would still show that left PFC function in attention increases when reasoning becomes more demanding relative to when it is less demanding. This is not an especially insightful claim about reasoning specifically, and it adds little additional new knowledge about the function

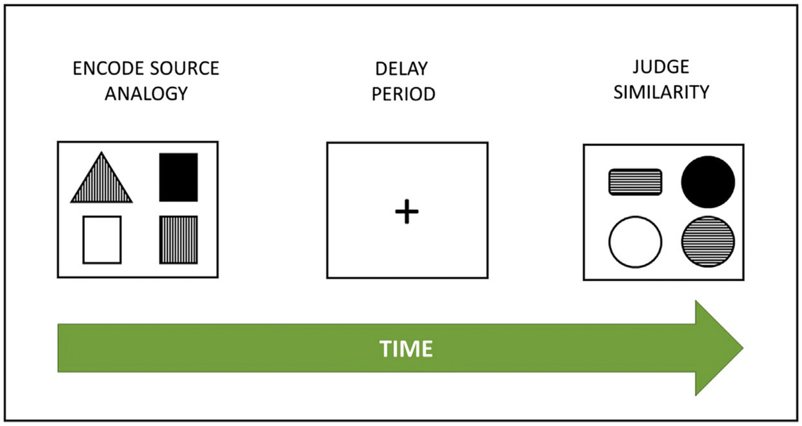


FIGURE 3.11 Wharton et al. (2000) conducted an early analogical reasoning study using pattern stimuli resembling items from matrix reasoning tasks.

of the left PFC. Another challenge of interpretation is that other regions of the brain are also involved in attention, and if the left PFC were in some way deactivated, perhaps these other regions (such as the right PFC) could stand in for the typically active left PFC. The activity in other regions could enable relational reasoning or analogical comparisons to be made all the same. Other methods were needed in order to more firmly establish a role of the left PFC in reasoning. Researchers needed additional methods to establish causal links that would enable a stronger conclusion than the idea that the left PFC is relevant to executive control regions and reasoning operations.

Transcranial magnetic stimulation (TMS) is a method that allows researchers to influence brain function in a safe and temporary way. The method is carried out by positioning a handheld magnetic coil that is roughly the size of a Ping Pong paddle over the scalp of a participant. Pulses of electrical current can be delivered through the coil. Electrical coils will induce magnetic fields when activated, so magnetic energy can be transmitted through the scalp and skull to induce electrical currents within the cortex. TMS pulses are relatively noninvasive. Participants may experience a tapping or tingling sensation around the location of the coil on the scalp, but the method is otherwise safe. TMS may be used to temporarily fatigue neurons in a given area of the brain thereby inducing a temporary reduction in function. Alternatively, TMS can be used in a facilitatory manner to induce an enhancement in brain function, though again this has only a temporary effect. Typically, researchers will deliver TMS while a participant performs a cognitive or perceptual task on a computer. Participant's heads are usually restrained using a metal frame with a chinrest, similar to those used in optometry offices. Many researchers will use MRI scans of the participants' brain in order to better target a particular area of the cortex. TMS is limited by its inability to effectively reach deep brain structures such as the amygdala and hippocampus, both of which are located within the interior portion of the temporal lobes.

Not long after the Wharton et al. (2000) study was published, Charles Wharton, Jordan Grafman, et al. conducted a TMS study of analogical reasoning with Babak Boroojerdi at the Cognitive Neuroscience Section of the National Institute of Neurological Disorders and Stroke. Boroojerdi et al. (2001) presented participants with the same type of analogical relation stimuli previously used in the Wharton et al. (2000) study. The analogies were presented in a sequential match-to-sample format in which a source picture was initially shown to the participants followed by the target picture, which had to be matched to the source. In an analogical condition, participants had to match the target to the sample based on the relations (color, pattern, or shape) among the four figures (see Fig. 3.11). In a literal comparison control condition, participants had to match the target to the sample based on perceptual similarity, thereby controlling for most aspects except relational reasoning. The task was also run with both target and sample pictures being presented simultaneously to allow for a comparison that reduced working memory demand on the participants. Facilitatory repetitive transcranial magnetic stimulation (rTMS) was delivered as participants performed all four conditions (sequential analogy, sequential literal match, simultaneous analogy, and simultaneous literal match). The rTMS was applied over the left PFC, as that region had been activated by analogical reasoning in the Wharton et al. (2000) PET study. Boroojerdi et al. also applied rTMS over the right PFC to evaluate its role in reasoning, as well as regions of the motor cortex as a control condition. Lastly, sham rTMS was simulated with the coil directed away from the participants' scalp, while these same task conditions as a procedural control condition.

Results of the study by Boroojerdi et al. (2001) further confirmed an important role for the left PFC in relational analogical reasoning. The application of rTMS over the left PFC resulted in faster performance for the analogy trials (by 100–200 ms) presented both sequentially and simultaneously. The speed increase did not occur during the sham conditions. No speed advantage occurred when rTMS was applied over the right PFC, suggesting that there was indeed a lateralization present for relational reasoning, as had been suggested by the earlier PET and fMRI studies. Results from the trials in which left motor cortex received rTMS were less clear, as speed advantage was also present for both simultaneous and sequential analogy conditions for the left motor rTMS condition, along with a speed advantage for the literal match condition when presented under sequential conditions (Fig. 3.12). This last finding with rTMS applied to the left motor cortex further supported the left frontal dominance in processing analogical relations, but also suggested that some of the speed decrease may be linked to a motor speed advantage rather than a purely cognitive processing speed advantage.

Up to this point in the literature there was a growing focus on the role of the PFC in reasoning. Over the next several years this focus gravitated toward the left PFC and then the left rostral PFC. There was some concern that the contributions of cognitive neuroscience were becoming limited to merely localizing all activity down to the functions of single brain areas. The focus on PFC localization was complicated by the fact that matrix reasoning and fluid intelligence tests are specifically designed to minimize (or exclude) the contributions of semantic memory or prior knowledge. It became clear that future cognitive neuroscience efforts would have to address the role of prior knowledge and how areas supporting semantic memory interacted with attention and working memory in order to facilitate human reasoning.

Integrating Relational and Semantic Information

Verbal analogies are a commonly used tool for investigating intelligence. Verbal analogies typically take the form of four-term "A is to B, as C is to D" comparisons. Such examples have been used in developmental investigations of children's knowledge and abilities.

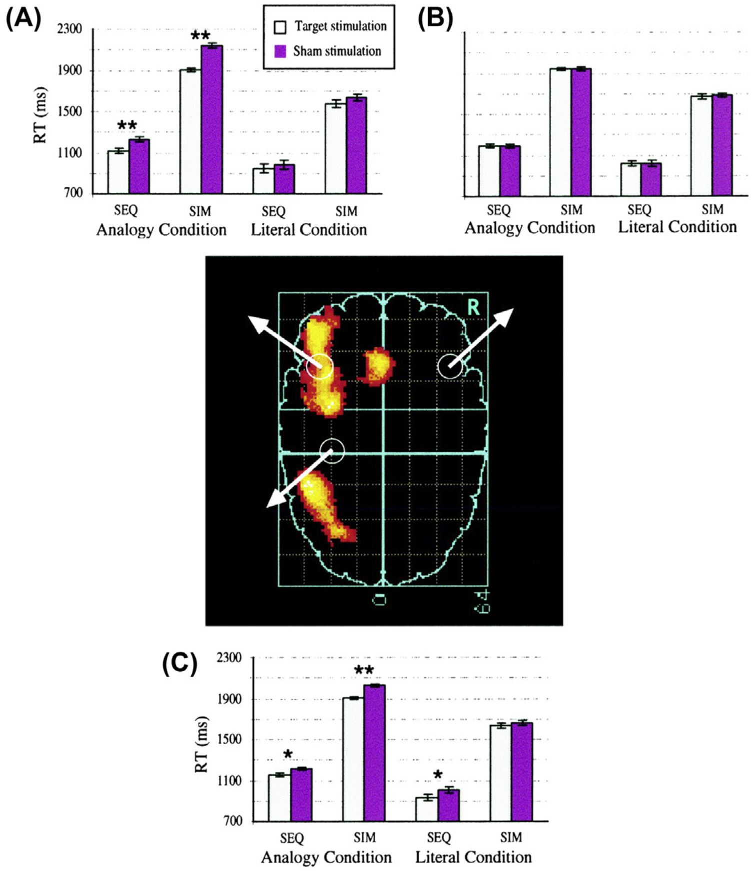


FIGURE 3.12 Boroojerdi et al. (2001) used transcranial magnetic stimulation (TMS) to induce facilitatory input over the left (A) and right PFC (B) and the left (C) and right motor cortex during a reasoning task. The results indicated that the left, but not right, PFC enhanced speed of processing during analogical reasoning. *Credit: Boroojerdi, B., Phipps, M., Kopylev, L., Wharton, C. M., Cohen, L. G., & Grafman, J. (February 27, 2001). Enhancing analogic reasoning with rTMS over the left prefrontal cortex, Neurology, 56(4), 526–528. http://dx.doi.org/10.1212/WNL.56.4.526:1526-632X.*

Robert Morrison et al. (2004) published a paper describing deficits in verbal four-term reasoning experienced by participants who had acquired frontal lobe damage through fronto-temporal lobar degeneration, a form of dementia. These individuals had difficulty solving for the fourth term. For example, participants were asked to complete the analogy "*Play* is to *Game*, as *Give* is to *?*" and the answer choices *Party* and *Take* were offered. Individuals with frontal damage had difficulties selecting *Party* as the correct answer, especially when the alternative, such as *Take*, was highly associated with the third term, *Give* in the current example. This finding suggested that in addition to relational processing, verbal or semantic relatedness was also a necessary skill in order to reason analogically. We will discuss this study further in Chapter 7 on disorders of reasoning.

Shortly after the Morrison et al. (2004) study appeared, Bunge, Wendelken, Badre, and Wagner (2005) published one of the first neuroimaging studies to investigate verbal analogical reasoning with four-term items. This study compared brain activation when participants evaluated analogical relationships compared to other situations when they simply evaluated how two words were related. For example, participants evaluated the analogical problem *Bouquet* is to *Flower*, as *Chain* is to *Link*. In this case, all relational information must be considered in order to evaluate whether this is a valid analogy. Meanwhile, other sets of four words were presented in a similar format such as *Note* and *Scale* followed by a second unrelated pair, such as *Rain* and *Drought*. Participants judged whether rain and drought were related. Findings indicated that the judgments of word relatedness involved the left inferior frontal gyrus, a region also known as the left ventro-lateral prefrontal cortex (VLPFC). Meanwhile, the comparison between analogies and evaluations of semantic-relatedness indicated that additional left PFC areas were preferentially active in response to the relational evaluations. These areas were more anterior to the VLPFC and supported the earlier relational integration studies that had been carried out using nonlinguistic materials. The possibility that different left hemisphere specialized areas of the frontal cortex supported word retrieval and relational integration was introduced.

Additional research on the neural correlates of verbal four-term analogical reasoning appeared with additional considerations about task demands and more specificity within the brain. Green, Fugelsang, Kraemer, Shamosh, and Dunbar (2006) asked participants to evaluate four-term verbal analogies and found that a left frontopolar cortical region (an area similar to the RLPFC) was active when participants had to evaluate the validity of an analogy, such as *Planet* is to *Sun*, as *Electron* is to *Nucleus*. This analogical evaluation condition was compared with a categorization condition, in which two word pairs had to simply be evaluated regarding the presence of independent semantic relations among the word pairs. For example, participants would indicate that a true categorical relationship was present between *Cow* and *Milk* and between *Duck* and *Water*, but there was no analogical relationship formed. The subtraction of semantic word-pair evaluations revealed only this single left frontopolar cortex location of activation associated with the evaluation of the relations between analogies. The results of this study largely converged with those of the study by Bunge et al. (2005) in the localization of a left PFC site of activation associated with relational comparisons in analogies that occurred over and above the activation associated with simpler semantic association evaluations.

Studies that include semantic knowledge, or word relatedness comparisons, along with relational processing may be more effective in simulating real-world reasoning. This may be especially the case when comparing the work on four-term analogies to the earlier Raven's progressive matrices fMRI studies, or those that involved evaluations of visual and spatial dimensions in simple analogical pairings. Both types of studies converged in isolating a left anterior PFC regional activation that escalated with relational processing demands. Only the semantic analogical reasoning tasks were capable of differentiating the left PFC areas that are associated with evaluations of semantic association from those that specifically process relational comparisons. While the work of Bunge et al. (2005) and Green et al. (2006) included more semantic information than several of the prior relational reasoning studies, it remained unclear to what degree other factors in semantic knowledge influenced the brain activity associated with relational comparisons.

Researchers had described the interesting properties of distant, or remote, analogies for several years prior to the emergence of functional neuroimaging studies. The term *remote analogical reminding* had been used to describe instances in life when two quite different situations could be seen as abstractly similar based on overall relationships that were analogous, despite great differences in the surface similarities of these situations (Gentner, 1983; Wharton, Holyoak, & Lange, 1996). An example can be found in fables. In the next chapter we will discuss one of Aesop's famous fables. In the fable, a thirsty crow confronts a pitcher of water in which the water level is too low to drink from. The crow solves the problem by adding several stones in order to raise the water level in the pitcher to the point where he can drink. This situation is analogous to a modern day person who finds herself unable to complete a marathon due to being out of shape. She may work persistently running on a daily basis and adding more stamina each time, until at last she can complete the race and experience a feeling of accomplishment. The relationships between the crow and the runner are apparent when one considers the system of relations among the elements. The crow matches to the runner, the quenching of thirst corresponds to the completion of the marathon, and the adding of stones matches the sequence of effort applied in daily running practice. Some within the research community even argued that the more remote an analogy was in terms of domains, the more interesting and satisfying the situations are to compare.

The next fMRI study from Green, Kraemer, Fugelsang, Gray, and Dunbar (2009) investigated the effect of remoteness, or distance, on the brain activity associated with analogical reasoning. These researchers were interested in capturing some elements of creativity and innovation that are characteristic of analogical reasoning. This is difficult to achieve using fMRI, as one must use numerous trials that are repetitive, short in duration, and relatively simple. A clever way to address the remoteness of relational comparisons is to consider word association values. Over the years cognitive and language research labs had quantified the association values between words. This can be done in a variety of ways by local co-occurrence in texts, or by having large numbers of participants generate associations to words. Green et al. used a technique called latent semantic analysis to calculate the similarity between the meanings of words within a high dimensional semantic space (Landauer & Dumais, 1997). The results of the latent semantic analysis enabled Green et al. to experimentally vary the distance among the analogies that were presented in their study using fMRI. As in the prior study (Green et al., 2006) four-term analogies were presented in which participants evaluated the validity of an analogy; however, in this study the distance of the analogical relationships was varied. In highly associated within-domain analogies, the word pairs presented were relatively close in semantic distance. For instance, in the analogy *Nose* is to *Scent*, as *Tongue* is to *Taste*, all of the items link strongly to parts of the face. In distant cross-domain analogies, comparisons were made between word pairs that had low association values, meaning that they do not tend to be strongly associated semantically. An example of such an analogy is *Nose* is to *Scent*, as *Antenna* is to *Signal*. This second example may appear to be a more clever analogy, as the distance highlights the relevant relation (sensory device and what it detects) more clearly when other non-relational semantic associations are absent. Green et al. interrogated the same left frontopolar PFC region that they had previously identified as being modulated by relational processing when compared to semantic association control conditions. Interestingly, the task-related activity within the left frontopolar PFC increased as the semantic distance among the words increased. This finding indicates that the rostral or frontopolar PFC is responsive to relational information, but that this area is also sensitive to the degree of remoteness between concepts being relationally compared.

The dual role of the left frontopolar PFC in relational integration and in managing remote domains is consistent with much of what has been known about the area from a cell physiology perspective. The PFC is considered to be heteromodal, meaning that this area of the brain is involved in numerous functions. The PFC is also considered to be highly plastic in that PFC cells change their response patterns rapidly in accordance with new task demands. The frontopolar PFC may be considered to be at the top of a hierarchy of information within the brain, serving as a coordinating or switching area to allocate attention toward particular goals or to integrate different types of information that may be recalled in the service of a reasoning task. To properly understand the role of the PFC in reasoning, key questions emerge that earlier functional neuroimaging studies had not yet begun to investigate. Does PFC activity vary depending on the type of reasoning task? What about the format of the task? When do different reasoning demands maximally engage the PFC? Such questions would begin to be addressed as additional labs became involved in the functional neuroscience of reasoning.

The Cognitive Operations of Reasoning

The studies we have reviewed thus far emphasized conditions that emphasized relational integration or managing multiple relations compared to ones that did not. The control conditions in those tasks largely focused on information complexity or visual complexity, but critically the controls did not emphasize any processing of relational or abstract information. A challenge with such studies is that the relational task conditions typically involved several cognitive abilities and those abilities may occur in a sequence over time that is not easily captured during neuroimaging tasks. This is due to the number of consistent and repetitive trials needed to average over blood oxygen signal to create a statistical map of brain activity. In studies conducted without the use of neuroimaging methods, there has often been a strong sense that information relevant to solving reasoning tasks becomes available at different points in time and a variety of cognitive abilities including perception, attention, working memory, and long-term memory are involved in reasoning tasks. These abilities are applied either to infer new knowledge about the world, make a prediction, verify a possibility, or overcome some sort of obstacle. Further investigation of cognitive abilities is a key to understanding the process of reasoning and how it is carried out within the brain. It is also important in this endeavor to specify when and why different cognitive abilities are utilized and for what purpose during the process of reasoning. This type of approach is necessary to understand the temporal dynamics of reasoning and ought to make it possible to avoid overly fixating on single brain areas as "modules" for reasoning. Most would argue that reasoning is too variable and diverse to be fully captured by a single module within the brain.

The act of inference is central to the purpose of making a relational or analogical comparison. Consider the example "Ink is to Quill, as Paint is to Paintbrush." In the analogical reasoning studies we have reviewed thus far in the chapter, participants were typically asked to evaluate whether or not there was an analogical relationship between those word pairs. Participants would likely have evaluated the relationship between the first two words, ink and quill, noting that the quill is used as a writing tool to transmit ink onto paper. The individual would then note that a paintbrush is used to spread paint purposefully onto other surfaces, thus there is a valid analogy present and that this should be verified in the experimental task. Given such instructions, the participants' goal is to simply state whether there is a relational match. This type of task is effective in eliciting rapid evaluations of analogies, which occur by the participant evaluating relations, comparing them, and proceeding to answer whether they are relationally similar. Such cognitive processing may explain why there is a high degree of similarity between many of the matrix relational reasoning tasks and those in which analogies were evaluated, as both tasks emphasize the same process of evaluating relations and then comparing them serially.

In everyday life, we make analogical inferences. These occur when we consider relational similarity between an older source analog from our past and a new situation that is unfamiliar and contain some unknown aspects. In this case, we may wish to use information from the source analog to infer a new conclusion about the new situation. For example, if you are familiar with driving cars, but have never ridden a motorcycle before, you may relate parts of the motorcycle to similar gadgets in your car. You might compare the motorcycle throttle to the gas pedal on your car and infer that it will make the vehicle accelerate in the same way that your car accelerates after you press the pedal. Note that simply comparing two word pairs to verify if they share a relational correspondence does not appear to capture much of the cognitive work that goes into the reasoning process in the motorcycle-car analogy. If we consider this new task goal of inferring new knowledge on the basis of a prior situation, it can change the cognitive operations a great deal and indeed the task itself becomes different in some key aspects.

In 2010 my own lab published a study that investigated four-term analogical reasoning as it occurred over time, with an eye toward understanding when frontal, parietal, and temporal lobe regions became active in the reasoning process. We were particularly interested in the inference stage of processing an analogy, which to that point had not been the subject of much direct investigation in neuroscience studies. A small change in the method from those prior four-term analogy neuroimaging tasks we have reviewed so far made this possible. In our experiment, we initially presented participants with the first pair of the analogy, for example, *spyglass* and *ship*. These objects were presented in picture form, which offered the participants an opportunity to consider the pair, visualize them, and relate them to one another. We next presented a picture of only one element of the second relational pair. For instance, participants saw an image of a *periscope*. We instructed our participants to imagine what would go with the periscope in order to complete the same relational association that had been illustrated by the spyglass and ship. Lastly, we presented either a picture of a valid item to complete the analogy, such as a *submarine*, or a picture of an unrelated item that did not satisfy the analogy, such as an *elephant*. Participants indicated with a button press whether or not the new item did indeed complete a valid analogy or not. We included two control conditions that were presented in an identical manner to the analogy condition, but with different instructions. In a semantic control, we simply asked participants to infer a semantic associate of the third item. For example, they could generate the concept of a sailor to go with the periscope, as this item did not need to share relational similarity with the source pair. In a perceptual condition, we simply asked participants to think of an item that looks similar to the third item in the analogy. For instance, they could think of a flashlight or a piece of drain pipe as looking similar to the periscope. Again, there was no need to consider similarity to the source pair in this condition. Refer to Fig. 3.13 for an example of the task design.

We gathered fMRI data from all phases of this task. This enabled us to ask a new set of questions: what areas of the brain become active during the evaluation of a source pair? What areas activate during the process of analogical inference? How are these areas of activity similar or different to the situation in which one evaluates a full analogy when the final item has been presented? By presenting analogies in this sequence and by spacing all of these task situations out in time, we hoped to better appreciate what different cognitive demands were associated to activity in what areas of the brain.

Let's reconsider the earlier example involving the pen-paintbrush analogy to better illustrate the difference in this analogical reasoning study compared to the prior ones. To infer something about the paintbrush that would complete a relational correspondence, we must initially appreciate that a quill is dipped in the ink with the goal of applying the ink to write or draw. The thinker is likely

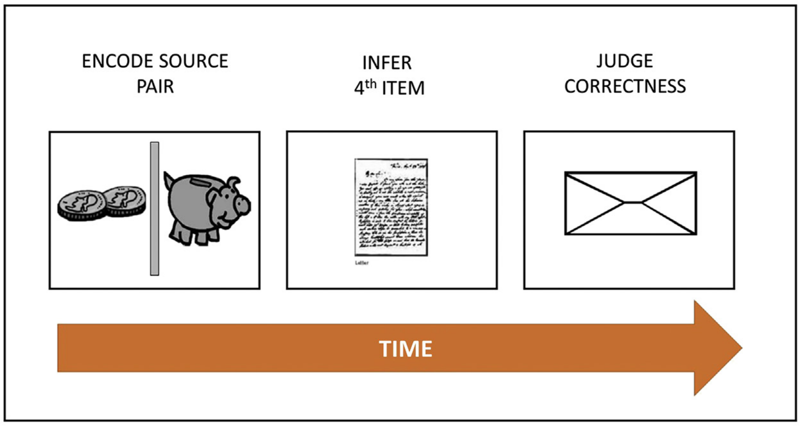


FIGURE 3.13 Krawczyk et al. (2010) evaluated four-term analogy problems as presented over time. This design allowed comparisons of the brain activation associated with encoding relational pairs, inferring a matching relational paired item, and verifying the solution.

to settle upon a dominant association that links these two items, in this case focusing on their functions. Next, the thinker is presented with the paintbrush, at which point he must perform a constrained memory search to isolate the function of the paintbrush and generate an item that is functionally equivalent to the role that the inkbottle plays in the earlier source analog. Lastly, when a paint can is presented, he or she must either note that it agrees with the generated item, or is capable of fulfilling the role of containing the paint that would be needed to allow the paintbrush to function. The cognitive processes present across the sequence are different at each point in the task and it is still much simpler than what we often do during the course of everyday analogical reasoning. To understand the brain processes that are evoked during different aspects of reasoning, it is necessary to go to the trouble of trying to isolate the different steps in a reasoning task and their associated cognitive processes. Only then can we begin to grapple with which functions are needed for reasoning and how these functions are correlated with brain activity.

The results of the Krawczyk, McClelland, Donovan, Tillman, and Maguire (2010) study helped to clarify the importance and functional significance of many of the brain areas noted in prior matrix reasoning and analogical relational reasoning studies. Fig. 3.14 shows a series of the major task-related areas of the frontal lobes: the left DLPFC, Left inferior frontal gyrus (LIFG), medial PFC, left middle frontal gyrus, posterior medial PFC, and right DLPFC. Note that several of these areas, especially the left DLPFC and LIFG, have already been discussed in association with evaluations of semantic relatedness in analogical reasoning tasks and in processing relational information. In addition to these areas, we found numerous other frontal areas including those on the medial surface of the PFC and on the right side as being strongly relevant in this analogical reasoning task. In each major area we plotted activation levels at the encoding period, in which the first two items in the analogy were evaluated, the mapping and inference period, in which a new item was generated in response to the presentation of the third picture, and a response period in which the participant indicated whether or not the final item shared analogical, semantic or perceptual similarity with the third item. We observed some very similar trends across all of the PFC areas involved in the task. The analogical reasoning condition evoked the greatest activity in the encoding period and in the inference period. Forming analogies evoked much more activity within all of the frontal areas except the right DLPFC during the encoding period. This makes sense, as participants knew that they would not need to consider that first pair during the semantic and perceptual conditions. At inference, the left DLPFC, LIFG, left MFG, and medial PFC all showed greatest activation levels for the analogical condition over the others. Several reverse trends were evident at the response period, when analogies typically evoked the lowest levels of activity relative to both the semantic and perceptual conditions. This was accompanied by slightly quicker and more accurate responses during the analogy condition. This pattern is consistent with the possibility that participants had more work to do during this last period evaluating a potentially new fourth item that they had not generated on their own previously for the semantic and perceptual conditions, while the analogical matches were more constrained and therefore easier to evaluate.

My colleagues and I followed up the Krawczyk et al. (2010) fMRI study with a second study that used event-related potential (ERP) recordings in order to better understand the temporal or timing dynamics of the cognitive processes that are involved in four-term analogical reasoning. ERP is a technique used to record electrical potentials off the scalp using electrodes placed on

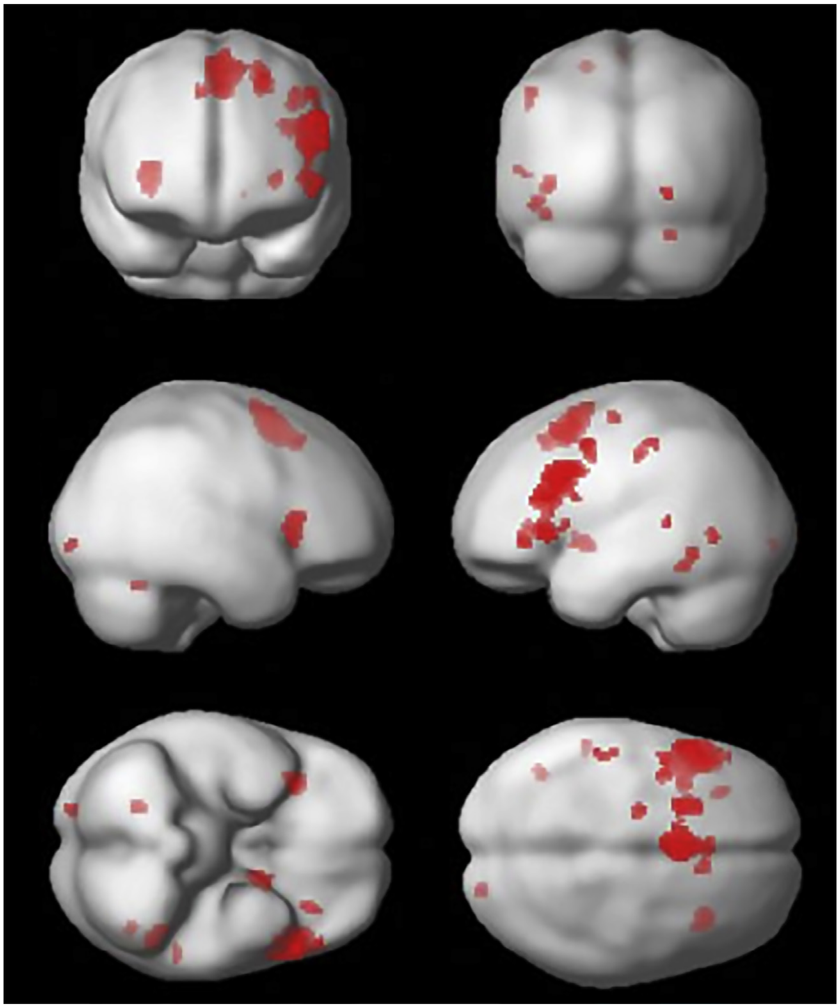


FIGURE 3.14 Results from the analogical reasoning study by Krawczyk et al. (2010) indicated strong left PFC involvement at the time of relational encoding relative to the other task periods. Inferring a new item that would complete a pair sharing relational similarity involved both left frontal and posterior areas and the right DLPFC.

the head. Typically, all of the electrodes are arranged in a mesh and positioned within a cap that enables recording from most of the head simultaneously. ERPs are derived from a recording technique termed electroencephalogram (EEG), which dates back to Hans Berger's pioneering work in the 1920s. Unlike fMRI, which depends upon changes in blood oxygen in order to generate its signal, EEG signals are generated directly from the electrical activity of the brain. Thus, ERPs are much better suited to answering questions about the timing of events within the brain relative to fMRI. The challenge with ERP data is that they cannot be directly traced to a particular brain region. This can be understood by considering the spatial layout of the EEG cap. Typical electrode placements occur centimeters apart and are simultaneously recording signal that is generated from many thousands of neurons that are active many millimeters below the scalp and are passing through the meninges (the outer coverings of the brain), the skull, and hair. What ERPs do very well is provide excellent timing information.

For our analogical reasoning study using ERPs (Maguire, McClelland, Donovan, Tillman, & Krawczyk, 2012), we used nearly an identical experimental setup for this experiment again comparing analogical inferences and verification to the generation and evaluation of semantic associations, as well as perceptually similar objects. During the encoding period, two images, A and B, were presented simultaneously (e.g., spyglass: ship) for 2 s. After a 1-s fixation period, the mapping and inference period included presentation of the C item for 3 s (e.g., periscope). After another 1-s fixation break, the response period included presentation of either the correct target D item (e.g., submarine) or a false item that did not complete the analogy. This period also lasted 2 s. We analyzed data collected from an EEG cap fitted with 64 electrodes while participants evaluated the analogy, semantic and perceptual problems. Data analysis used a Principle Components Analysis to address the locations relevant to analogical encoding and those associated with mapping and inference processes. Additionally, we wished to harness the ERP methods' superior temporal resolution to better understand the timing of these task events in analogical reasoning. Results indicated that left frontal electrodes showed the strongest response to the analogical encoding period. In this case a significant positive EEG waveform was present for encoding of analogical relations relative to the semantic condition beginning 500 milliseconds (ms) after the stimuli were presented and continuing through 1200 ms after stimulus presentation. Meanwhile, a different pair of spatial and temporal patterns were detected supporting the mapping and inference stage of the task. In this case, the left frontal electrodes yielded a negative response, 400–600 ms, post-stimulus that was significantly different for analogies over the perceptual condition. A second site of activity was from the posterior electrodes and was a positive waveform, 700–900 ms, post-stimulus onset.

When considered together the fMRI study by Krawczyk et al. (2010) and the ERP study by Maguire et al. (2012) tell a more complete story than one would achieve with either method alone and reveal some striking features about the neural basis for analogical reasoning. First, the encoding of a relational pair preferentially engaged several left-lateralized PFC regions including the DLPFC, IFG, and middle frontal gyrus in the fMRI task over the other conditions. Additional encoding related areas included the medial PFC and right DLPFC. The ERP study indicated a left frontal electrode site, which contributed a negative waveform, 500–1200 ms, into the encoding period. When considered together, we found that the left PFC is important for the evaluation of relational information in analogical comparisons and that it may also be involved in the maintenance of relational information, but additional PFC regions are also important for the encoding of relational information and these include the right dorsal and midline areas of the PFC. Secondly, the mapping and inference phase of the tasks indicated again that the left PFC was important and potentially somewhat dominant over the right PFC. The positive ERP waveform for the encoding period followed by the negative waveform for mapping and inference are consistent with different cognitive processes driving these effects.

An overall understanding of brain responses evoked by a reasoning task requires a variety of methods that can be applied to the task and an appreciation of the fact that several different cognitive processes are likely to be engaged in any reasoning task. In the domain of analogical reasoning, the initial results from fMRI indicated a left-sided dominance and an almost modular focus on the rostrolateral or frontopolar regions as being critical for relational integration or comparison. Later studies indicated that semantic distance was also a factor that modulated the left frontopolar cortex along with relational integration. When considered in isolation, the encoding of a relational pair is supported by numerous PFC regions with a left-sided dominance overall. The process of inferring a new item that would complete a pair sharing relational similarity involved both left frontal and posterior areas and also showed some degree of modulation in the right DLPFC.

Note that in nearly all of the studies we have reviewed in this section involved four-term analogies. Reasoning in everyday life is much more fluid, less repetitive, and may involve more diverse and complex material. We will discuss these differences in future chapters that delve further into the types of challenges that we actually reason about in everyday life. It is worth remembering that neuroscience studies must be precise about timing and certainty of what task the participant is engaging in at a given moment. Only under those conditions can researchers capture meaningful differences associating cognitive events to activity within the brain.

FUNCTIONAL NEUROANATOMY OF KNOWLEDGE

The Temporal Lobes and Networks for Concepts

To this point in the chapter we have focused to a large degree on the role of the PFC in reasoning. While findings suggest that the PFC is probably the most relevant brain region for functions such as integrating information, dealing with novel situations, and managing our working memory and emotions, it is not the only area important for reasoning. There are many situations in everyday life in which our reasoning does not heavily involve integration or working memory.

Consider a situation in which you are on a date at a nice restaurant. You need to balance your ordering, manners, timing of actions, and track and contribute to the conversations. Having a successful date can be considered a reasoning problem. The date could go poorly for a variety of reasons, but it will surely go disastrously if you are unable to organize your behavior in the restaurant setting. The way to approach this challenge is to use your prior knowledge and avoid overly relying on working memory to get you through the date. You will almost surely do worse in your attempts to appear smooth and sophisticated if you are unfamiliar with restaurant dining, wine, menus, and general protocol. We call this type of knowledge a *schema*, which is a connected set of background information that will apply most of the time. Schemas are developed by observing and being part of many instances of an event or situation. In the case of a restaurant schema, we are likely to have a general sense of what takes place at most restaurants on most date settings. You will expect to arrive on time and possibly have to wait if you have not made a reservation. You will expect to be seated and provided with water and menus. You will expect to have time to study the menu and will have to balance this activity with tracking and keeping up your part of the conversation. You will also expect the food to arrive within a reasonable timeframe and be given an opportunity to consider dessert before the bill is brought. These features of restaurants occur in a particular order and schemas that have this property are often called *scripts*. The reasoning problem of carrying off a successful date, or at least not a disastrous date, probably relies much more on our schemas and scripts than it does on working memory and multitasking. Reliance on the schema will help free up your other mental abilities which you will likely need in order to focus on your partner, observe their needs and actions, and carry on a relevant conversation. You will also need to free your attentional resources for the all-important act of listening. The more you have to use your valuable attention and working memory skills to simply track and predict the mechanics and timing of the meal, the worse your date will likely go.

Schemas and scripts rely mostly upon the cortex of the temporal lobes. The idea of schemas dates back to the work of Sir Frederic Bartlett, a pioneering British experimental psychologist. One of Bartlett's classic works was a study on the "War of the Ghosts," a story about Native American tribal warfare (1932). This story is provided in Box 3.2. The "War of the Ghosts" included elements common to many tribal battle situations, such as rowing a canoe, shooting arrows, and invading a village. The story also included a variety of unusual elements, such as mystical occurrences; being wounded, but feeling no pain; and the notion that the battle had occurred among ghosts. Bartlett asked people to retell the story at a later time, which ranged from just minutes to years later. Findings indicated that when people were asked later on about the story, they would tend to recall a distorted version of the story and the distortion increased along with the passage of time since the original reading of the events. While people will retell the story in slightly different ways, there appear to be some consistent principles by which we distort memories for stories like this.

Frequently people will remember the core elements during a retelling of the "War of the Ghosts," while failing to include many of the specific details. In other words, people tend to recall a gist of the story. Examples include the idea that there were two young men; they went down river, a battle occurred and one was shot by an arrow. The specific details of the story; however, are often lost in the retelling. People are much less likely to accurately recall the name of the village (Egulac), what the young men were initially hunting (seals), or what the young men did when they heard war-cries (hid behind a log). Over time, people would also begin to merge their retelling of the story with their own background knowledge. In other words, they recalled those details that were consistent with their schemas about tribal warfare. Along with this merging of information, people tended to leave out details that were not schema-consistent. As time passed, a greater proportion of the reported story elements were schema-consistent, and many fewer unusual details were reported. At the latest points in time the story was also edited quite a lot suggesting that many details either become lost or simply are less accessible over time.

The work of Bartlett (1932) on the "War of the Ghosts" recall tells us quite a bit about the reconstructive nature of memory and this in turn is informative about how we reason based on schemas, expectations, and gist (Fig. 3.15). First, the changes in recall of the story tell us that memory for details changes over time. This observation is consistent with the notion that memories are reconstructed each time we recall them and represents a fundamental point about how we see the world. The presence or absence of cues is critical to determining whether we will recall a memory. Some cues may evoke memory for certain details, while other cues will evoke other details. Second, the retelling of a story indicates that we tend to build up a general sense of what was important to us about the situation over time. Rarely do we need to extensively focus on all aspects of a narrative or a situation. Our own version of the memory will vary from others, as observed in the "War of the Ghosts" retelling. This is shaped initially by the limitations of perception and attention determining how much information we can perceive at any given moment. The specific aspects of a situation that we attend to will further delineate what we will recall later. Given that both perception and attention are limited, we are already bound to lose many details of any occurrence. Lastly, the specific cues we happen to encode that associate with a memory further shape what we will recall later.

The act of recalling a memory is also heavily influenced by internal factors, and this brings us back to the concept of schemas and scripts. While we have relatively less control of what we can actually see, hear, touch, or smell at any given moment, we have a voluntary ability to attend to some things and not others. This has been compared to a signal-to-noise problem, in which some aspects of a situation will be amplified over other aspects and accompanying this amplification is often an elevation of brain activity within our cortex. The fact that we are able to move our focus of attention is another reason that some of our memories will be different than others. What guides our attention may often be what is relevant to us and this is dependent upon our background knowledge to a large

BOX 3.2

THE "WAR OF THE GHOSTS"

One night two young men from Egulac went down to the river to hunt seals and while they were there it became foggy and calm. Then they heard war-cries, and they thought: "Maybe this is a war-party.". They escaped to the shore, and hid behind a log. Now canoes came up, and they heard the noise of paddles, and saw one canoe coming up to them. There were five men in the canoe, and they said:

"What do you think? We wish to take you along. We are going up the river to make war on the people."

One of the young men said, ""I have no arrows."

"Arrows are in the canoe," they said.

"I will not go along. I might be killed. My relatives do not know where I have gone. But you," he said, turning to the other, "may go with them."

So one of the young men went, but the other returned home.

And the warriors went on up the river to a town on the other side of Kalama. The people came down to the water and they began to fight, and many were killed. But presently the young man heard one of the warriors say, "Quick, let us go home: that Indian has been hit." Now he thought: "Oh, they are ghosts." He did not feel sick, but they said he had been shot.

So the canoes went back to Egulac and the young man went ashore to his house and made a fire. And he met everybody and said: "Behold I accompanied the ghosts, and we went to fight. Many of our fellows were killed, and many of those who attacked us were killed. They said I was hit, and I did not feel sick."

He told it all, and then he became quiet. When the sun rose he fell down. Something black came out of his mouth. His face became contorted. The people jumped up and cried.

He was dead.

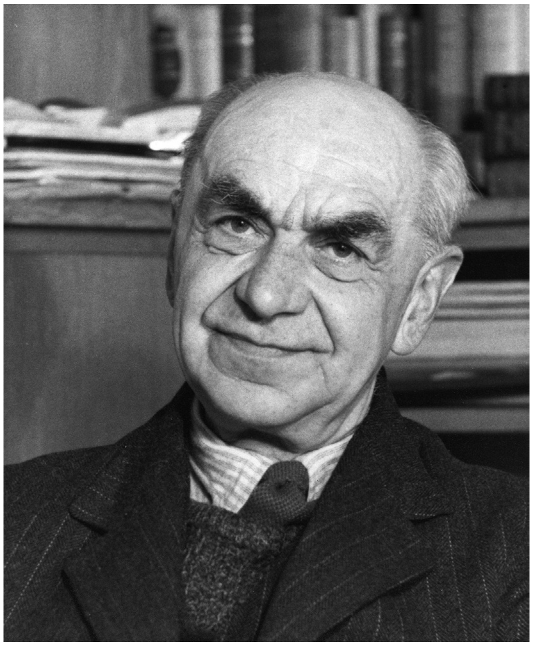


FIGURE 3.15 Sir Frederick Bartlett (1886–1969) is well-remembered for his pioneering work on schemas in semantic memory. *Courtesy of the Department of Psychology Archive, University of Cambridge.*

degree. This is perhaps most apparent when we develop a working model of something with our minds and then proceed to test our assumptions by reading more information, talking to others, or seeking out new experiences that will inform us. The influence of both attention and semantic memory can be seen from this example. Our reasoning processes involve language and symbolic coding of information in our networks of semantic memory.

DEDUCTION, INDUCTION, AND THE BRAIN

A Strong Philosophical Distinction

Deductive and inductive reasoning are often compared and contrasted. Fundamentally, deductive reasoning guarantees the validity of a conclusion, provided that the premises are true. For example, if the premises state that all planets are round and that the Earth is a planet, it is valid to conclude that the Earth is round. Induction is an argument from a specific example toward a general rule. Induction is different than deduction, as inductive reasoning does not guarantee a valid conclusion, only a probable conclusion based on experience. For example, if we have heard that both the brown recluse spider and the black widow spider are poisonous, we may be tempted to conclude that all spiders are poisonous and should be avoided. This could be true or possibly false, based on our degree of experience with spiders. We will consider this distinction in greater detail in Chapter 9.

Philosophers have noted the strong distinction between deductive and inductive reasoning. Aristotle wrote about the distinction between inductive and deductive arguments, emphasizing the power of deduction primarily. Later, David Hume emphasized the shortcomings of inductive reasoning, as it was prone to errors due to the need to infer a global rule based on the experience of only a subset of instances. In purely academic terms, there is a strong distinction between deductive and inductive inferences. This can translate into very real consequences in everyday life when we may make errors due to inductive inferences, while deductive logic will always give us a valid conclusion, provided the premises are valid.

Evaluating the differences between these styles of reasoning has been a long-standing goal of psychologists, who are often most interested in the cognitive processes that occur during inductive inference and how these may differ from those engaged during deductive inference. Adding further clarity between these processes at a biological level became possible due to the emergence of functional imaging techniques, notably fMRI and EEG. These techniques have also offered an additional method to investigate the various conditions that are possible within deduction or induction, such as comparisons of deductive reasoning with realistic content compared to pure deduction based on novel and arbitrary rules.

Evidence From Neuroimaging

Some of the first studies to be conducted comparing deductive and inductive inference used the PET methodology discussed earlier in this chapter. Goel, Gold, Kapur, and Houle (1997) published a PET imaging study with the intriguing line "The seats of reason?" in the title. In this study, Goel et al. asked participants to solve classic deductive reasoning problems such as "All men are mortal, Socrates is a man; therefore, Socrates is mortal." In response to these deductive statements, participants were asked to indicate whether the first two statements (all men are mortal and Socrates is a man) entailed the third statement (Socrates is mortal). Participants were also asked to evaluate inductive statements such as: "Socrates is a cat; Socrates has 32 teeth; therefore, all cats have 32 teeth." In the inductive inference condition participants were prompted to indicate whether the third statement was plausible given the first two. Both deductive and inductive reasoning activated areas of the occipital, frontal, and temporal cortex. Goel et al. performed a subtraction analysis, in which the inductive and deductive reasoning conditions were directly compared. Deductive reasoning trials were subtracted (or contrasted) with those involved in inductive reasoning, and this analysis revealed one area more strongly activated by induction, the left superior frontal gyrus on the inside or medial wall of the left frontal lobe. This finding indicated that a potential difference between induction and deduction was associated with the functions of the left PFC. Note that this is a different region than had been observed in other relational reasoning studies around this same period, such as those of Wharton et al. (2000) and Duncan et al. (2000). The region that Goel et al. had identified was more medial than those other studies. It should be noted that there was not a processing load manipulation in the Goel study, as deductive and inductive reasoning about verbal statements are more difficult to vary with different levels of complexity or processing load in the manner that can be done for processing relational aspects of geometric figures discussed earlier in the chapter.

Goel, Gold, Kapur, and Houle (1998) conducted a second PET imaging study, in which they sought to better understand the possible links between spatial cognition and deductive reasoning. In this study, participants solved three types of deductive reasoning tasks (categorical syllogisms, three-term spatial items, and three-term nonspatial relational items) while they underwent PET imaging. All of these deduction conditions showed similar activation patterns when compared to a baseline condition that simply required people to judge the meaning of different statements. In all three cases, deductive reasoning led to activation of a large section of the left lateral PFC in the middle frontal gyrus. Additional areas of activation included the inferior and superior temporal gyri and the left cingulate cortex. Similar regions of the left PFC and temporal cortex appear to be relevant in other types of reasoning problems, notably the solution of analogical and relational reasoning problems.

The content of the problems can dramatically influence deductive reasoning. Goel et al. next conducted an fMRI study of deductive reasoning in which they manipulated the level of realistic content contained within deductive reasoning problems. All deductive syllogisms take a form that can be generically expressed as statements lacking semantic content. For example, these can be stated as "All A's are B's, all B's are C's; therefore all A's are C's." People tend to reason differently about these conditional syllogisms than they might about realistic deductive arguments that contain semantic information. Deductive syllogisms that are content-rich can take the form "All lions are cats, all cats are mammals; therefore, all lions are mammals." Such statements may be faster and easier to process for people if the content is consistent with what they already know. If however, the content goes against what is already known, evaluating these kinds of statements can actually prove to be more difficult for people. Imagine a syllogism such as the following: "All lions are bears, all bears are lizards; therefore all lions are lizards." While this last statement is deductively valid, it does not accurately fit people's background knowledge about categorizing animals. In the fMRI study by Goel, Buchel, Frith, and Dolan (2000), participants evaluated deductive arguments lacking content ("all P are B, all B are C; all P are C"), congruent syllogisms with content that fit with people's knowledge ("all poodles are pets, all pets have names; all poodles have names"), and incongruent syllogisms that were at odds with people's knowledge ("all poodles are pets, all pets are vicious; all poodles are vicious"). In this study, activation for the deductive reasoning conditions over baseline revealed strong bilateral activation in bilateral fusiform gyrus, the left parietal cortex, left middle temporal gyrus, bilateral inferior frontal gyrus, and the basal ganglia bilaterally. These regions are broadly consistent with the earlier PET study by Goel et al. (1998), but fMRI revealed greater right frontal, occipital, and basal ganglia activation that had not previously been detected using PET. This difference may be related to the superior spatial and temporal resolution of fMRI over PET imaging for conducting functional studies. Syllogisms with and without semantic content yielded broadly similar patterns of activation with some differences. Notably, a conjunction analysis was used to identify areas of activation common to both conditions. These included the left inferior PFC, bilateral fusiform gyri, bilateral basal ganglia, and the right cerebellum. There were differences observed with subtractions of the content minus no content conditions leading to activation of the left temporal cortex, left inferior PFC, and bilateral occipital cortex. The reverse subtraction of no content conditions minus content conditions showed greater activation within bilateral occipital, parietal, frontal, and precentral gyrus. This suggests that there are some key processing differences that can be detected within several cortical areas when people reason about information they know about compared to when they evaluate deductive logic in purely formal and abstract terms. Lastly, Goel et al. reported greater activity within two bilateral PFC regions when participants evaluated syllogisms congruent with their world knowledge relative to ones that went against what they know.

These initial studies were successful at identifying the major sites of activation that are associated with deductive and inductive reasoning. They also provided some clues regarding the possible differences between formats of deductive reasoning and the differences between inductive and deductive inferences. Unlike the early studies of relational reasoning that we covered in earlier sections within this chapter, these syllogistic reasoning studies evoked broader and somewhat less-specific activation across the bilateral frontal lobes, along with other brain areas including the occipital cortex, and sometimes parietal and temporal regions. The cognitive operations underlying deductive and inductive forms of reasoning may be broadly similar and commonly involve working memory, semantic memory, and decision-making processes.

In summary, deductive and inductive inference appeared to show largely similar activation of several cortical areas. Among these regions is a left-lateralized set of areas involved in reasoning about real-world situations and content along with a right-lateralized focus for general deduction about arbitrary premises. Other research groups had similarly reported frontal, parietal, and occipital involvement in deductive reasoning about three-term series problems presented as syllogisms (Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002) and for solving deductive reasoning problems requiring the evaluation of conditionals and relational arguments, which is consistent with the engagement of visuo-spatial processes in reasoning (Knauff, et al., 2002). Fangmeier, Knauff, Ruff, and Sloutsky (2006) reported largely bilateral activation for a visuo-spatial deduction task, which indicated the possibility that the right-lateralized activity that had been observed in prior studies was potentially less clear due to visual and spatial processing. More work was needed to further specify which cognitive processes were linked to which brain areas in inductive and deductive conditions.

A Focus on Process-Related Areas in Deduction and Induction

Further research led to further discoveries about the core cognitive mechanisms of deductive and inductive reasoning. This later set of studies was also able to further map cognitive mechanisms onto the emerging brain activation data that were increasingly being collected. To further understand the relevance of the evoked activation in frontal, temporal, and occipital areas and across hemispheres, it became necessary to carry out experiments that enabled more precise mappings between specific points in the reasoning process and areas of activation. This approach is similar to the process investigators took to more precisely understand stages of analogical reasoning and the activity elicited in those tasks. Diana Rodriguez-Moreno and Joy Hirsch (2009) published the results of a multistage fMRI study that helped to clarify the likely associations between brain activity and specific parts of the deductive process. Participants in this study saw an instruction at the start of each trial indicating whether it was a deductive reasoning trial or a control trial that simply required participants to verify if a particular word appeared in the problem. This control condition enabled the researchers to estimate activation levels that were associated with text processing, attention, and judgment that were not specific to deductive reasoning specifically. After the instruction cue, participants were shown a premise either as text on a screen or auditorily through headphones in order to better understand the role of language processing modality on brain activation. An example premise would be "All politicians recycle glass bottles." Next a second statement was presented, for example, "People who recycle glass bottles like wildlife." A short delay period followed this statement enabling the researchers to separate activation

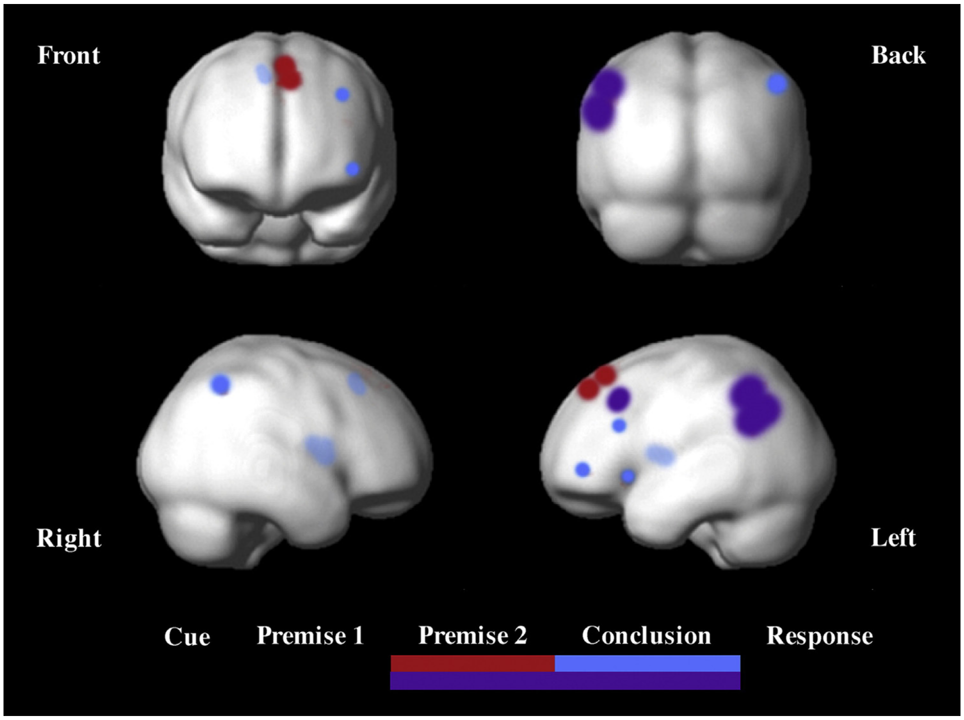


FIGURE 3.16 Rodriguez-Moreno and Hirsch (2009) reported a shift of activation from superior frontal and parietal regions during the second premise of a deductive reasoning task (presented in red) to areas of the left frontal lobe, bilateral parietal and caudate nucleus during the conclusion (presented in blue). Some of these areas showed sustained activity through the second sentence and conclusion (presented in purple). *From Rodriguez-Moreno, D., & Hirsch, J. (March 2009). The dynamics of deductive reasoning: An fMRI investigation, Neuropsychologia, 47(4), 949–961.*

from the premises to that evoked by the conclusion. Conclusions were presented as a third statement such as "All politicians like wildlife." Lastly, there was a word presented for use in the control condition. The detail of this design enabled a variety of important comparisons that helped to isolate the areas of activation that are specific to the reasoning process, while excluding other active areas that are related to other cognitive or sensory processes.

Rodriguez-Moreno and Hirsch (2009) reported a set of areas that were active in response to deductive reasoning demands when presented through auditory or written format. These included the DLPFC, VLPFC, frontopolar PFC, supplementary motor areas, posterior parietal cortex, inferior parietal lobules, and the caudate nuclei. Several of these same regions have been previously discussed in association with relational reasoning and with induction. In addition to the identification of areas involved in reasoning, Rodriguez-Moreno and Hirsch (2009) were able to specify at what points during the reasoning process brain areas became engaged. In the case of the task instruction cue and the first premise there were no differences for deduction than for the control task requiring verification of whether a word had been present in the premises. Additionally, no differences were observed between deduction and the control task for the first premise. Only at the point when the second premise was presented did the deductive and control conditions diverge, with deduction being associated with increases in activation within the left PFC, left parietal cortex, and angular gyrus. The authors described this period of the deductive task as being related to premise integration and conclusion generation. Similarly, the deduction task differed from the control task during the presentation of the conclusions at the point when participants would have had to evaluate the validity of the full syllogism. This comparison revealed activation within the left middle and superior frontal gyri, the right medial frontal gyrus, bilateral parietal cortex, and bilateral caudate bilaterally. Note that left frontal regions were activated during both the integration of premises and the evaluation of a conclusion, as was the parietal cortex (Fig. 3.16). Lastly, Rodriguez-Moreno and Hirsch's study of deduction compared the differences that format had in evoked brain activation. When presented in an auditory format, greater numbers of active voxels emerged for deduction relative to the presentation by text. An additional observation was that while similar frontal, parietal, and caudate areas were active from the period after the presentation of the second premise until the conclusion, the auditory condition tended to have a slightly delayed time course of activation.

The shift in research was moving toward investigation of stages in reasoning. The stages in deduction were also tested by Fangmeier et al. (2006), who evaluated simpler deductive conditions represented by letters. In this task, single letters were presented to participants in the following manner:

Premise 1: V X
Premise 2: X Z
Conclusion: V Z

Participants were able to evaluate these premises by considering them in terms of the sentences "V is to the left of X," "X is to the left of Z," and asked to evaluate whether the conclusion "V is to the left of Z" is valid. In this case, the comparison yields a valid conclusion. Fangmeier et al. were also able to evaluate the stages of deduction in this paradigm. The presentation of the first premise was limited to activation within temporal and occipital cortex, possibly linked to recognition and memory. The second premise activated the anterior PFC in a manner consistent with the prior results that we have discussed on relational reasoning in which integration of information is frequently marked by greater levels of activation within the frontopolar PFC. The final conclusion evaluation, or validation phase, was associated with greater activation within the PFC and lateral parietal areas, similar to those results reported by Rodriguez-Moreno and Hirsch (2009).

While the specific locations of activation differ, many of these same frontal and parietal regions were found to be involved in the stages of analogical reasoning reported by Krawczyk et al. (2010), which included the integration of two relations and the generation of a term to complete the problem. Both types of reasoning involved the encoding of two items or premises, integration of information across time, and the generation of a conclusion followed by verifying whether a final piece of information fit with the problem solution generated.

Work by Cocchi et al. (2014) tested this relationship directly by varying relational complexity conditions within a test of deductive reasoning. The task was a variation of the classic Wason card selection task (Wason, 1966) in which individuals are presented with a deductive reasoning rule such as "If a card has a vowel on one side, then it must have an even number on the other side." When people solve the Wason task, they are presented with numbers and letters and have to determine whether the card would need to be flipped over to see if the number or letter on the back side of the card could disprove the rule presented. For example, if a card has a letter A on one side, it would need to be flipped to see if it indeed had an even number on the back side. If it does not, then this would violate the rule. In Cocchi et al. experiment they also introduced a relational complexity variable such that some cards required greater amounts of relational information that would be necessary to determine whether the card flip would disprove the premise.

Functional MRI results from Cocchi et al. (2014) largely replicated those of Fangmeier et al. (2006) and those of Rodriguez-Moreno and Hirsch (2009) indicating strong DLPFC activation along with parietal activation among other areas in the premise evaluation phase, followed by DLPFC, frontopolar PFC, striatum, thalamus, and insula activity during the card decision phases, when deductive reasoning was maximally engaged. Additionally, Cocchi et al. reported increases in these PFC regions when relational complexity was increased. In this study, we see a pattern of convergence in which deduction and relational processing are considered as being governed by some of the same brain regions. This type of study indicates that the field had begun to move past investigations in which different types of reasoning were always predicted to engage different neural systems. It remains an open question to what degree there is overlap among the neural systems supporting different types of reasoning. Accomplishing this objective will require further research. Some clues to the resolution of the overlap are presented next in the following section on neural networks.

NEURAL NETWORKS OF REASONING

A Shift Toward Neural Networks

To this point in the chapter we have been focusing largely on the locations of activity associated with different aspects of reasoning and how these can be related back to cognitive events. Many of the neuroimaging studies of relational reasoning have been focused primarily on identifying which areas are most essential for discriminating between purely relational comparisons and those that involve semantic or perceptual similarity. Such comparisons lead us to conclude that the most critical area for relational reasoning is likely the left anterior PFC, which has been called either the rostrolateral PFC (Wendelken, Nakhabenko, Donohue, Carter, & Bunge, 2008) or the frontopolar PFC (Green et al., 2006). This regional focus should not be interpreted to mean that the left anterior PFC is a module purely for reasoning, or that it functions in isolation by carrying out the full extent of reasoning functions and passing that information on to other regions. If this were the case, then damage to the left anterior PFC could be predicted to lead to a catastrophic inability to perform relational reasoning tasks. This is not the case, as we will discuss in Chapter 7 covering disorders of reasoning. Frontotemporal lobar degeneration (FTLD) is a form of dementia that typically causes frontal lobe atrophy and loss of function. Individuals who develop frontal-variant FTLD often lose functionality in the anterior-most regions of the PFC first and progressively lose more frontal cortical function as the disease progresses. These patients do not typically reach a neurological clinic until their behavior becomes sufficiently disorganized that it cannot be ignored or attributed to simple changes in personality or lifestyle. At such a time, most patients will have lost dramatically more cortical function than just the anterior PFC, and some patients who are actively being treated for frontal-variant FTLD do not express a complete deficit in analogical reasoning, but rather differences in speed or reductions in accuracy (Morrison et al., 2004).

To understand reasoning in the brain, we must consider the myriad of brain regions that have been discussed as being activated specifically in association with both relational reasoning and semantic processing. Additionally, we must consider the connectivity within white matter that enables these brain regions to connect and share information.

A very useful technique has emerged called functional connectivity. This technique was originally carried out in resting-state studies. A resting-state fMRI scan consists of scanning a person while they are awake and thinking to themselves internally, rather than performing a task. Some resting-state scans require participants to keep their eyes open and perhaps focus on a single point on a screen, while other investigators have simply let participants think internally. What emerges from resting-state scans is a consistent feature of neural systems. Areas will tend to correlate with one another, which is an indication that neurons within these areas are likely communicating as their blood-oxygen demands rise and fall together. This finding was originally reported by Biswal, Yetkin, Haughton, and Hyde (1995). This group observed strong correlations between the right and left motor cortices. The areas of the brain that are actively correlated or connected are not random. There is a characteristic pattern for resting-state activity in a set of areas that are clustered toward the midline of the brain between the two hemispheres. These midline areas have been called the default mode network (Gusnard & Raichle, 2001), which includes areas such as the medial PFC and precuneus, a region in the middle of the parietal lobe. Other rest-state functional connectivity studies have identified other sets of commonly active neural regions that show spontaneous and persistent correlations. Among those most relevant to reasoning functions is the task-network or fronto-parietal control network, which is frequently engaged when someone exercises his or her working memory or cognitive control abilities. Similarly the cingulo-opercular network, including regions within the cingulate cortex and the frontal operculum, is thought to play a supporting role in directing attention and sustaining task focus.

Neural network approaches have been shown to be highly informative for reasoning studies providing a clear way to understand how cortical areas interact, rather than promoting the further search for isolated cognitive modules. Task-based connectivity was carried out assessing the correlations between areas within the deductive reasoning Wason card task carried out by Cocchi et al. (2014). In this study, the task-related set of areas, including the frontopolar PFC, DLPFC, anterior cingulate cortex, parietal and occipital lobes, and the anterior insula, was active during the deductive reasoning conditions. Cocchi et al. (2014) analyzed the correlations among these areas finding that there was greater functional connectivity among the areas when relational complexity increased. The areas that were active during this deduction task overlapped with the fronto-parietal control network and the cingulo-opercular network that we had discussed earlier. Both of these networks are strongly active during cognitive control tasks. Cocchi et al. (2014) suggested that the enhanced connectivity of these network areas may indicate that the DLPFC and frontopolar PFC serve as important nodes in the neural networks that carry out reasoning functions. Further work will be needed to better delineate whether there are regions that do act in a more modular way despite the widespread connectivity.

SUMMARY

The series of studies on the neuroscience of reasoning covered in this chapter suggest a variety of summary points. These studies also point to several unresolved questions that will guide future research directions.

Species that possess large amounts of cortex and especially association cortex tend to show advanced reasoning skills. These species include primates, dolphins, and whales, as well as elephants and certain birds. The relationship of large brains to complex behaviors may be driven by whether an organism is a predator. Predators must possess a wide array of behaviors in order to locate and subdue prey.

Many of the early fMRI studies of reasoning suggested that the PFC was particularly important. Additionally, we have noted that the frontopolar PFC and left DLPFC have important roles based on imaging and electrophysiological studies. Some network studies indicate that the frontopolar PFC has been reported as an active component of numerous reasoning tasks across several domains, including visuo-spatial relational reasoning, analogical relational reasoning, deductive and inductive reasoning. This indicates that the PFC may serve a coordinating or control function, possibly integrating wide-scale activity across the brain, rather than operating as a relational module. We have also noted the impact of materials on reasoning studies in the brain. Materials invoke our semantic memories, which are supported by temporal lobe regions. The integrated sets of information that we can use in reasoning are called schemas or scripts. Brain network interconnectivity looks to be an especially promising area toward capturing the complexity of neural processing in reasoning and may further clarify some of the roles of specific brain areas that have been linked to reasoning in various forms.

END-OF-CHAPTER THOUGHT QUESTIONS

1. A large amount of cortex is associated with advanced reasoning skills. Can you think of some exceptions to this tendency?
2. The relationship of large brains to complex behaviors may be driven by whether an organism is a predator. What specific behaviors of predators may have driven their brains to develop?
3. Functional neuroimaging provided the ability to map brain regions that support reasoning. What are some of the benefits and limitations of this type of study?
4. Many of the early neuroimaging studies of reasoning suggested that the prefrontal cortex was particularly important. Is this area developed for reasoning or for other cognitive processes, such as memory and attention?
5. Neural network studies indicate that the frontal lobes may serve a coordinating or control function. How might we test this possibility?
6. The materials used in an experiment have a large impact on the results of reasoning studies. Materials invoke our semantic memories, which are supported by temporal lobe regions. Can you think of some ways to test this?
7. The integrated sets of information that we can use in reasoning are called schemas or scripts. How could you adapt the script or schema concept into a testable neuroimaging study?
8. Brain network interconnectivity looks to be an especially promising area for reasoning. Are all types of reasoning supported by networks, or do some types appear to rely upon regional brain activity?

References

Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance Medicine*, *34*, 537–541.

Boroojerdi, B., Phipps, M., Kopylev, L., Wharton, C. M., Cohen, L. G., & Grafman, J. (February 27, 2001). Enhancing analogic reasoning with rTMS over the left prefrontal cortex. *Neurology*, *56*(4), 526–528. http://dx.doi.org/10.1212/WNL.56.4.526:1526-632X.

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*, 239–249.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., et al. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage*, *14*, 1136–1149.

Cocchi, L., Halford, G. S., Zalesky, A., Harding, I. H., Ramm, B. J., Cutmore, T., et al. (2014). Complexity in relational processing predicts changes in functional brain network dynamics. *Cerebral Cortex*, *24*, 2283–2296. http://dx.doi.org/10.1093/cercor/bht075. Epub April 5, 2013.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A neural basis for general intelligence. *Science*, *289*, 457–460.

Fangmeier, T., Knauff, M., Ruff, C. C., & Sloutsky, V. (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience*, *18*, 320–334. http://dx.doi.org/10.1162/jocn.2006.18.3.320.

Fuster, J. M. (2006). The cognit: A network model of cortical representation. *International Journal of Psychophysiology*, *60*, 125–132.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

Goel, V., Buchel, C., Frith, C., & Dolan, R.J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage*, *12*(5), 504–514.

Goel, V., Gold, B., Kapur, S., & Houle, S. (1997). The seats of reason: A localization study of deductive & inductive reasoning using PET (O15) blood flow technique. *NeuroReport*, *8*(5), 1305–1310.

Goel, V., Gold, B., Kapur, S., & Houle, S. (1998). Neuroanatomical correlates of human reasoning. *Journal of Cognitive Neuroscience*, *10*(3), 293–302.

Green, A. E., Fugelsang, J. A., Kraemer, D. J., Shamosh, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research*, *1096*, 125–137.

Green, A. E., Kraemer, D. J., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2009). Connecting long distance: Semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex*, *20*, 70–76. http://dx.doi.org/10.1093/cercor/bhp081.

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience*, *2*, 685–694.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*, 253–270. http://dx.doi.org/10.1037/h0023816.

Knauff, M., Mulack, T., Kassubek, J., Salih, H. R., & Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research*, *13*, 203–212.

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.

Kroger, J. K., Saab, 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*, 477–485.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Maguire, M. J., McClelland, M. M., Donovan, C. E., Tillman, G. D., & Krawczyk, D. C. (2012). Tracking cognitive stages in analogical reasoning with event-related potentials. *Journal of Experimental Psychology: Memory, Learning and Cognition*, *38*, 273–281.

Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L., et al. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience*, *16*, 260–271.

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

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence, 1938, sets A, B, C, D, and E*. London: H. K. Lewis.

Raven, J. C. (1960). *Standard progressive matrices: Sets A, B, C, D, & E*. London: H.K. Lewis & Co.

Robin, N., & Holyoak, K. J. (1995). Relational complexity and the functions of prefrontal cortex. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 987–997). Cambridge, MA: MIT Press.

Rodriguez-Moreno, D., & Hirsch, J. (March 2009). The dynamics of deductive reasoning: An fMRI investigation. *Neuropsychologia*, *47*(4), 949–961. PMID:18835284.

Rumelhart, D., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge: MIT Press.

Sternberg, R. J. (2000). The holey grail of general intelligence. *Science*, *289*, 399–401.

Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth: Penguin Books.

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*, 682–693.

Wharton, C. M., Grafman, J., Flitman, S. S., Hansen, E. K., Brauner, J., Marks, A., et al. (2000). Toward neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology*, *40*, 173–197.

Wharton, C. M., Holyoak, K. J., & Lange, T. E. (1996). Remote analogical reminding. *Memory & Cognition*, *24*, 629–643.

Further Reading

Bartlett, F. C. (1920). Some experiments on the reproduction of folk stories. *Folk-lore*, *31*, 30–47.

Ferrer, E., O'Hare, E. D., & Bunge, S. A. (2009). Fluid reasoning and the developing brain. *Frontiers in Human Neuroscience*, *3*, 46–51.