Reasoning and Problem Solving: Models

A K Barbey and L W Barsalou, Emory University, Atlanta, GA, USA

Introduction

Neuroscience research has increasingly provided evidence that informs theories of reasoning. We begin by defining classic forms of reasoning; and summarizing psychological contemporary theories. We then review empirical evidence from neuroscience that bears on these theories. Finally, we summarize current challenges and identify promising areas for future research.

Definition and Types of Reasoning

Reasoning is a hallmark of human thought, supporting the process of discovery that leads from what is known or hypothesized, to what is unknown or implicit in one's thinking. Reasoning can take the form of deductive inference, whereby the evidence guarantees the truth of the conclusion. Alternatively, when reasoning depends on conditions of uncertainty, it takes the form of inductive inference, whereby the evidence provides only limited support for the truth of the conclusion. The deductive inference that a storm has emerged is supported by evidence that it is raining, whereas observing lightning clouds provides only limited support for the inductive inference that it will rain. Many forms of reasoning typically depend on conditions of uncertainty, including problem solving, causal reasoning, and analogical inference.

Problem solving refers broadly to the inferential steps that lead from a given state of affairs to a desired goal state. For example, deciding who to vote for in a presidential election, diagnosing a patient on the basis of observed symptoms, or preparing for a mountain climbing expedition all depend on problem solving. Problem solving often requires the process of planning, namely, formulating a method for attaining a desired goal state. Planning a mountain climbing expedition, for example, requires scheduling the season of the expedition by predicting the consequences of taking the trip during different seasons. This is accomplished by modeling the situation and observing the consequences of possible actions (e.g., the occurrence of snow in winter introduces new challenges).

Prediction and explanation further depend on causal reasoning (i.e., the ability to infer causal relations). For example, knowledge that cold winter weather can cause an accumulation of snow in the mountains, which in turn can cause an avalanche, supports the prediction that an avalanche is possible during winter.

Finally, new predictions and explanations can be generated by analogical reasoning, whereby relations from one domain are mapped onto another domain. For example, students can conceptualize particle motion in an atom through an analogy to the solar system.

Theoretical Perspectives on Reasoning in the Psychological Literature

The psychological literature contains various theories that address the cognitive architecture and symbolic representations that underlie reasoning. After reviewing these theories, we turn to neuroscience evidence that bears on them.

Cognitive Architecture

One major distinction can be drawn between theories that view the mind as containing specialized reasoning modules, versus theories that view the mind as containing general-purpose reasoning systems. According to the modular view, the mind consists of specialized modules that are unavailable to conscious awareness and deliberate control (cognitively impenetrable), and that are only able to process specific types of information (information encapsulation). Advocates of this view have proposed a diversity of modules that underlie reasoning, including modules for semantic inference, communicative pragmatics, social exchange, intuitive numbers, spatial relations, naive physics, and biomechanical motion. If cognitive architecture emerges from an underlying neural architecture, then strong modular views predict that the neural systems for reasoning should be relatively localized, implementing modules that are cognitively impenetrable and informationally encapsulated.

Alternatively, dual-process theorists propose that reasoning is based on two general-purpose systems: an associative system and a rule-based system. The associative system uses basic cognitive operations such as association, similarity, and memory retrieval to produce primitive judgments quickly and unconsciously. The rule-based system reflects more evolutionarily advanced mechanisms that implement reasoning procedures deliberately and consciously. For example, inductive reasoning largely depends on the retrieval and evaluation of world knowledge, whereas deductive reasoning depends on rule-based, formal procedures.

The dual-process theory further motivates the cognitive demand hypothesis: when people have little time and limited processing resources, incentives, and/or external aids available for making a judgment, they only use the associative system when reasoning. Conversely, when people have more time and greater processing resources, incentives, and/or external aids available, they also use the rule-based system. The dual-process model predicts that reasoning recruits different neural systems, depending on cognitive demand. When tasks are easy, the associative system is sufficient for correct reasoning and primarily recruits left inferior frontal gyrus, especially Broca's area, in addition to the temporal lobes and posterior parietal association cortex. As tasks become more difficult, the rule-based system becomes required for correct reasoning and recruits the prefrontal cortex, especially the ventrolateral subregion, which has been implicated in rule maintenance.

A second theory that provides a general-purpose account of reasoning is grounded in a somewhat different pair of systems: the linguistic system and the conceptual system. According to this view, the brain's language system initially produces relatively superficial information about a reasoning problem, such as word associates, syntactic structures, etc. As linguistic forms become generated, their meanings become increasingly represented in the conceptual system. A key assumption of this approach is that superficial processing based on linguistic representations may often be sufficient for adequate reasoning performance. When it is not sufficient, conceptual representations must be generated to produce more sophisticated reasoning. Thus, consistent with the dual-process model, this view predicts that differences in the cognitive demands of a reasoning task (i.e., superficial vs. deep processing) will differentially engage neural systems.

In general, both dual coding frameworks predict that reasoning will recruit neural systems that support two forms of coding. One important set of systems underlies language processing, including the left frontotemporal language system. A second set of important systems underlies conceptual processing, mental simulation, and imagery, including bilateral sensorimotor areas.

Representation in Reasoning

A second major distinction can be drawn between theories of reasoning that are based on amodal versus modality-specific representations. According to standard theories, the modality-specific states that are active while perceiving an entity are redescribed into amodal representations that bear no correspondence to the neural systems producing them. For example, the concept 'chair' is represented by redescribing the modality-specific states that underlie the perception of a chair into amodal representations (e.g., feature lists, semantic nets), processed by classic symbolic mechanisms (e.g., predication, argument binding). Thus, amodal theories predict that conceptual processing will recruit brain regions outside sensorimotor areas that underlie language and rules, including left prefrontal and superior temporal regions implicated in formal, rule-based operations.

In contrast, embodied theories of knowledge propose that knowledge and meaning are grounded in modality-specific representations. Increasing empirical evidence from both the cognitive and the neuroscience literatures suggests that modality-specific representations underlie higher level cognition. According to this framework, concepts are represented by simulating the modality-specific states that were initially activated during perception, action, and interoception. Embodied theories propose that simulations are organized at a higher level by simulators that integrate information across a category's instances. Over time, for example, visual information about how cakes look becomes integrated in a 'cake' simulator, along with gustatory information about how cakes taste, somatosensory information about how they feel, motor programs for interacting with them, emotional responses to experiencing them, and so forth. The result is a distributed system throughout the brain's modality-specific areas that establishes conceptual content for the general category of 'cake.' Thus, embodied theories predict that conceptual processing will recruit a broadly distributed system of modality-specific brain regions.

Embodied theories further motivate the task specificity hypothesis. According to this hypothesis, the neural areas underlying a particular type of reasoning, such as deduction, may show little in common, as the specific materials and tasks vary. Because different materials and tasks produce different patterns of modality-specific activation, the same general form of reasoning does not show a single, stable pattern, and no common areas may emerge. In contrast to the 'cognitive demand hypothesis,' this view predicts that significant differences in the task and/or materials – even those that do not affect cognitive demand – are likely to dominate neural activation more than the type of reasoning performed.

Mental Models versus Mental Logic

A third major distinction can be drawn between theories of reasoning that are based on visuospatial models versus theories that are based on logical operations. Mental-model theory proposes that deductive and inductive reasoning depend on spatially organized mental models. According to this view, an argument is evaluated by generating alternative models of its premises, where each model represents possible circumstances that render the premises true. If all the models constructed from the premises are consistent with the truth of the conclusion, then the argument is judged to be deductive or valid. If, however, only a limited number of models render the conclusion true, then the argument is judged to be inductive or probabilistic.

Given the proposed role of mental models in both deductive and inductive reasoning, this theory predicts that these forms of reasoning will recruit common neural systems. Specifically, mental-model theorists have predicted that deductive and inductive reasoning will both primarily recruit right-hemisphere regions. Furthermore, to the extent that mental models represent situations spatially, this view predicts that parietal and occipital regions implicated in visuospatial processing will be engaged.

In contrast, mental logic theory offers an account of deductive reasoning that is based on the application of formal deductive rules according to formal syntactic operations. Thus, rule theorists have predicted that reasoning is likely to recruit left prefrontal and superior temporal regions implicated in formal, rule-based operations.

Empirical Evidence on Reasoning from the Neuroscience Literature

We next present empirical evidence from neuroscience that bears on the psychological theories just reviewed. We begin with research on deductive reasoning, and also address research that compares deductive and inductive reasoning. We then review research on problem solving, causal reasoning, and analogical reasoning.

Deductive Reasoning

The brain systems that implement deductive reasoning depend on whether the reasoning problem consists of familiar versus unfamiliar semantic content. When reasoning about familiar semantic content (e.g., All dogs are pets / All poodles are dogs / Therefore, all poodles are pets), a left frontotemporal system is engaged, including left inferior frontal cortex (BA 47), left middle/superior temporal cortex (BA 21/22), and left temporal pole (BA 21/38). Previous research has implicated this system, not only in deductive reasoning, but also in memory and language tasks that employ familiar semantic content. In general, linguistic processing appears central to all these tasks.

In contrast, reasoning about unfamiliar semantic content (e.g., All P are B / All C are P / Therefore all C are B) activates a bilateral frontoparietal system, including bilateral dorsal (BA 6) and inferior (BA 44) frontal lobes, bilateral superior and inferior parietal lobes (BA 7), and bilateral occipital lobes (BA 19). This pattern of activation is also found during the processing of spatial information, and is similar to the neural activity observed while people make transitive inferences about geometrical shapes.

Thus, deductive reasoning recruits left and right prefrontal cortex asymmetrically as a function of familiarity. Across both familiar and unfamiliar deduction problems, left prefrontal cortex is generally active, suggesting that this region is necessary for deductive inference. Conversely, right prefrontal cortex is engaged only when a problem involves unfamiliar semantic content or a conclusion that conflicts with prior beliefs (e.g., No harmful substances are natural / All poisons are natural / Therefore, no poisons are harmful). Whereas language may often dominate familiar reasoning, language and spatial/visual processing may often be central for unfamiliar reasoning.

The brain systems that implement deductive reasoning also depend on whether a reasoning problem produces correct versus incorrect conclusions. Research on this issue has used inhibitory belief problems, namely, problems whereby individuals must inhibit a highly accessible belief that could interfere with correct reasoning (e.g., No addictive things are inexpensive / Some cigarettes are expensive / Therefore, some cigarettes are not addictive). Drawing a correct conclusion on these problems requires that individuals (1) detect the conflict between their prior beliefs and the logical inference, (2) inhibit the prepotent response associated with their belief bias, and (3) engage the appropriate reasoning mechanisms. In contrast, drawing an incorrect conclusion on these problems results from failing to detect the conflict between beliefs and logical inference, and/or failing to inhibit the prepotent response associated with a belief bias.

When people draw correct conclusions on inhibitory belief problems, right inferior prefrontal cortex becomes active. When they draw incorrect conclusions, ventromedial prefrontal cortex is active instead. Activation in right inferior prefrontal cortex when drawing correct conclusions appears to reflect the detection and/or resolution of the conflict between belief and logic. Conversely, activation in ventromedial prefrontal cortex when drawing incorrect conclusions appears to reflect the role of nonlogical mechanisms, perhaps associated with greater affective processing.

In summary, the neural systems that underlie deduction vary considerably, depending on task factors and cognitive demand. Consistent with the task specificity hypothesis, the areas that support deduction vary with the familiarity of the materials, and with whether belief violations occur and are detected. Consistent with the cognitive demand hypothesis, more neural areas are recruited for difficult unfamiliar problems than for easier familiar ones. Consistent with the modularity view, left prefrontal cortex generally appears active across most deduction paradigms, suggesting that it is essential for deductive inference.

Deductive versus Inductive Reasoning

Many experiments have asked individuals to perform both deduction and induction in the same experiment, so that the neural circuits underlying these two types of reasoning could be distinguished. One neuroimaging experiment used a categorical syllogism task consisting of three conditions – deduction, induction, and baseline – to assess this issue. In the deduction condition, individuals received a valid or invalid categorical syllogism (e.g., None of the bakers play chess / Some of the chess players listen to opera / Therefore, some of the opera listeners are not bakers). The task was to indicate whether the conclusion was valid (i.e., whether the conclusion was guaranteed by the truth of the premises). In the inductive reasoning condition, individuals only received invalid categorical syllogisms (e.g., Some of the computer programmers play the piano / No one who plays the piano watches soccer matches / Therefore, some computer programmers watch soccer matches). The task was to indicate whether the conclusion was probable or improbable. The deduction and induction conditions were each compared to a baseline condition, in which a categorical syllogism with anomalous semantic content was presented (e.g., All the engineers own a computer / None of the engineers has been to school / Therefore, all the people who own computers are married). Importantly, the categorical syllogisms were fully counterbalanced across individuals so that the same materials occurred in every condition.

Deductive reasoning (deduction minus baseline) produced activation in the left dorsolateral frontal cortex (BA 6), broadly consistent with the left frontal activation observed by other researchers for familiar semantic content. When deduction was compared directly to induction, however, a different pattern emerged, namely, activation in bilateral posterior regions, with a right-hemisphere prevalence, including associative visual cortex (e.g., cuneus, precuneus, middle and superior occipital gyri), as well as right superior parietal lobe (BA 7) and thalamus. These areas have been reported for visuospatial tasks that require form discrimination and imaginative operations. These areas have also been reported for deductive reasoning tasks that employ visuospatial materials. Such activations suggest that people use visuospatial representations, such as Venn diagrams or Euler circles, to support deductive reasoning on categorical syllogisms. In addition, activation was also found in right anterior cingulate (BA 24/32), implicating attention and executive control in deductive reasoning. Inductive reasoning (induction minus deduction) revealed activation in left dorsolateral frontal (BA 8 and 10) and right insular cortices. These regions are known to be involved in probabilistic reasoning tasks that require the estimation of relative frequencies and other quantities (e.g., How fast do race horses gallop?).

A subsequent neuroimaging study extended the methods and design of the study just described to a deduction task that employed conditional statements instead of categorical syllogisms (e.g., If he is an electrician, then he spent two years in night school / He is an electrician and owns a computer / Therefore, he spent two years in high school). In contrast to the categorical syllogisms used in the previous study, visuospatial processing did not seem relevant to these conditional syllogisms. Thus, these researchers predicted that the new task and materials would not recruit brain regions that perform visuospatial processing. Consistent with this prediction, deductive reasoning (deduction minus induction) revealed a major focus of activation in right inferior frontal cortex (BA 44), in right anterior cingulate (BA 24), and in right middle temporal cortex (BA 21). These researchers concluded that this frontotemporal system constitutes a logic-specific network in the right hemisphere that is comparable to the language-specific network in the left hemisphere. Specifically, these researchers argued that this right-hemisphere system implements a calculus of mental transformations that underlie formal deduction.

In a further comparison, inductive reasoning (induction minus deduction) revealed large intense activations in the left inferior frontal (BA 47) and left insular cortices, in addition to left posterior cingulate (BA 31), parahippocampal (BA 36), left medial temporal (BA 35), and superior and medial prefrontal cortex (BA 9). These areas are broadly consistent with the left frontal (BA 8 and 10) and insular areas found in the previous study, known to be involved in the recall and evaluation of familiar world knowledge. Consistent with standard theories of inductive reasoning, the recall and evaluation of familiar world knowledge appear to play central roles during induction.

Another neuroimaging study developed a category learning task to assess the component processes of inductive reasoning. Of primary interest were the neural systems that support rule application versus rule inference during category learning. Stimuli consisting of novel animals were presented to individuals, who were asked to judge whether all the animals in a set were from the same category. In the rule application condition, a rule was provided that specified the criteria for category membership. In the rule inference condition, individuals had to infer the rule with no instruction. Each condition was further divided into an easy and difficult condition based on the computational demands of the task. These researchers found that rule inference (rule inference minus rule application) preferentially recruited bilateral hippocampus, an area in which activation is modulated by stimulus novelty. In contrast, rule application (rule application minus rule inference) revealed activation in the presupplementary motor area (BA 8). This area is implicated in the anticipation of motor activity and likely reflects an anticipatory response to category exemplars, a response that is absent when the categorization rule is unknown.

In addition to the episodic encoding of novel stimuli, inductive reasoning requires the generation and testing of hypotheses. For example, inferring the basis for category membership requires generating and testing possible rules, such as "has spots on the abdomen" or "has only two appendages" for a category of fictional animals. To assess the neural systems that underlie hypothesis selection, these researchers evaluated the task by difficulty interaction: {hard rule induction minus hard rule application} minus {easy rule induction minus easy rule application}. This comparison assessed the effects of increased difficulty due to subtle variations in the stimulus features of animals that satisfied or violated the category membership rule. The result was an activation in right lateral orbital prefrontal cortex (BA 47 and BA 11), an area implicated in complex reasoning tasks such as analogical and metaphorical transfer. Across different types of reasoning, difficult problems often activate this area.

In summary, current findings from neuroimaging studies of deduction and induction again suggest that the neural bases of reasoning are highly sensitive to the particular tasks and materials employed. For example, deductive reasoning from conditional statements engages a right frontotemporal system, thought to support the application of formal deduction rules, whereas deductive reasoning from categorical syllogisms activates a parieto-occipital system that supports visuospatial processing. Furthermore, the latter result differs from the findings of other researchers, who have found that a left frontotemporal system supports deductive inference in a categorical syllogism task. The observed differences on the same fundamental reasoning process – deduction – probably reflect methodological differences between the tasks and materials used across studies. Furthermore, the fact that different materials are often used on deduction and induction tasks suggests that different activations for these two forms of reasoning may often be driven by differences in materials rather than by differences in reasoning.

Consistent with the task specificity hypothesis, the neural processes that underlie a given type of reasoning (e.g., deduction) may depend more on the tasks and materials used than on the type of reasoning *per se*. Although left prefrontal cortex is often active on deduction tasks (as described in the previous section), it was not active for some of the studies discussed in this section. Perhaps the one consistent finding so far is that reasoning about familiar materials tends to utilize left-hemisphere language and knowledge networks. Conversely, reasoning about less familiar problems tends to utilize bilateral systems that include right-hemisphere mechanisms. This pattern is consistent with theories that postulate two different reasoning systems, one that processes language, and one that processes spatial/visual information.

Problem Solving and Planning

Depending on whether a problem-solving task is well structured or ill structured, different brain systems are engaged. On well-structured problems, such as the Tower of Hanoi, the starting state, the goal state, and possible transformations are specified completely. For example, the starting state consists of three pegs mounted on a platform and three disks of varying sizes stacked in descending order on the first peg. The goal state is to stack the disks in descending order on the third peg. The possible transformations are restricted to moving disks such that (1) only one can be moved at a time, (2) any disk that is not removed must remain on a peg, and (3) a larger disk cannot be placed on a smaller disk. In such tasks, well-structured planning typically recruits left prefrontal cortex, including frontopolar, dorsolateral, and ventrolateral regions.

In contrast, an ill-structured planning problem is specified incompletely. In the Multiple Errands Task, for example, individuals are taken to an unfamiliar neighborhood and asked to complete errands, such as buying a loaf of bread, according to the following rules: you are to spend as little money as possible (within reason) and take as little time as possible (without rushing excessively). You are not to use anything not bought on the street (other than a watch) to assist you. You may perform task steps in any order. In such tasks, ill-structured planning typically activates right dorsolateral prefrontal cortex

(DLPFC), in contrast to the left prefrontal regions that support well-structured planning.

The neural systems underlying planning are further differentiated into subsystems for plan formation versus plan execution. One finding is that DLPFC supports sequential operations during plan formation, whereas medial ventral prefrontal cortex (MVPFC) plays a motivational role in plan execution. More specifically, DLPFC appears to support the generation of hypotheses and the construction of plan steps, whereas the MVPFC appears to support the affective processing required for plan execution (e.g., initiative and determination).

Lesion studies complement the neuroimaging studies just reviewed. One representative study evaluated problem solving in patients with frontal lobe lesions (FLLs) on the Water Jug problem. Patients had to construct unique action sequences that transferred specific quantities of water between jugs of different sizes (e.g., Jar A = 8 units of water, Jar B = 5 units of water, Jar C = 3 units of water). Specifically, individuals had to construct an action sequence that achieved a goal state given by the experimenter (e.g., Jars A and B must contain 4 units of water, and Jar C must contain 0 units). The researchers found that patients with FLLs struggled to make counterintuitive moves that were required to solve the task but that appeared to deviate from the desired goal state. Furthermore, left and bilateral FLL patients were more impaired than right FLL patients were, suggesting that poor performance in the Water Jug task was primarily linked to left DLPFC damage.

In summary, these findings demonstrate that problem solving and planning generally depend on prefrontal systems. Because DLPFC appears important for most problem-solving tasks, modular views receive support. Differences in task conditions modulate the active areas of the prefrontal cortex, however, providing support for the task specificity hypothesis. For example, well-structured versus ill-structured problems rely on left versus right prefrontal regions, respectively. This pattern is also consistent with theories that assume two different systems, such as dual-process and dual-code theories, support reasoning. Finally, distributed theories receive support, given that multiple systems typically underlie problem solving. For example, problem solving requires both planning and execution, and also both reasoning and motivation, with different neural systems supporting each component process.

Causal Reasoning

Several studies have evaluated the neural systems that support the perception of mechanical causation in the classic Michotte launching event. In a launching event, a ball travels horizontally across a computer screen and collides with a ball located in the center. The collision results in the second ball 'launching' away from the first, horizontally, across the screen, thereby eliciting the perception that the first ball caused the second to move. One study compared the neural response produced by the launching event to the response elicited by a control event in which the first ball passed below the second ball without a collision (noncausal condition). Of primary interest were the neural systems engaged when individuals judged either (1) the presence or absence of causation versus (2) the direction of the ball's motion. A reliable increase in medial frontal activation occurred for judgments of causality relative to judgments of ball movement. Moreover, this increase occurred during both the causal and noncausal conditions, suggesting that the signal increase was specifically associated with the process of making a causal judgment, not with the perception of actual causality.

Another study evaluated whether causal perception and causal inference rely on common or distinct hemispheric regions. Two callosotomy (split-brain) patients and a group of neurologically intact patients were tested. Of primary interest was neural activity in the left versus right hemispheres during (1) the perception of causal events (i.e., the Michotte launching event) and (2) causal inference tasks when the relation between a candidate cause and an observed effect had to be inferred (rather than perceived directly). Perception of causality and causal inference depended on different hemispheres of the divided brain. Whereas causal perception engaged the right hemisphere, causal inference engaged the left-hemisphere.

Another study assessed the brain systems for processing evidence that was either consistent or inconsistent with an individual's existing causal beliefs. Individuals received evidence on the effectiveness of drugs designed to relieve depressive symptoms. Two factors were manipulated: the plausibility of the theory that explained the drug's action, and the consistency between theory and data. When individuals reasoned with evidence that was consistent with existing causal beliefs, a network of brain regions widely associated with learning and memory was engaged, including the caudate and the parahippocampal gyrus. In contrast, when they reasoned with inconsistent evidence, a different pattern of activation occurred that is widely associated with error detection and conflict resolution, including the anterior cingulate cortex (BA 24/32), posterior cingulate, and precuneus (BA 7). The researchers concluded that people's beliefs and expectations act as a filter during evidence evaluation. When evidence is consistent with existing causal beliefs, the neural systems underlying those beliefs implement causal reasoning. When evidence is inconsistent with existing beliefs, a different neural system detects this inconsistency and triggers the construction of a novel causal explanation.

The observed findings demonstrate that causal reasoning is supported by a broadly distributed neural system that is highly sensitive to the causal reasoning task (i.e., causal perception versus causal inference) and the consistency of causal evidence with existing beliefs. These findings provide support for theories of reasoning that are (1) based on distributed rather than localized representations (e.g., dual-process, dual-code, embodied theories), (2) incorporate the role of existing knowledge rather than operating on the basis of purely logical representations (e.g., dual-process, dual-code, embodied, mental-model theories), and (3) advocate the task specificity hypothesis (e.g., dual-process, dual-code, embodied theories).

Analogical Reasoning

Several studies have found that analogical reasoning engages frontopolar cortex. Furthermore, different components of analogical reasoning appear to differentially engage frontopolar versus dorsolateral prefrontal areas. Whereas dorsolateral areas are recruited for processing externally generated information (e.g., the monitoring and manipulation of presented facts), frontopolar areas are recruited additionally for the evaluation and manipulation of internally generated information.

One study assessed the neural systems that support analogical reasoning in the Raven's Progressive Matrices task. Study participants received a 3 × 3 matrix of figures with the bottom right figure missing, and had to infer the missing figure by selecting one of four possible alternatives. Participants received three types of problems that differed in their degree of relational complexity (0-relational, 1-relational, 2-relational). The 0-relational problems involved no relation of change and thus required no relational processing. The 1-relational problems involved one relation of change in either the horizontal or vertical dimension, and thus required relational reasoning. Finally, the 2-relational problems involved two relations of change, in both the horizontal and vertical directions, and thus required even more relational reasoning.

A region-of-interest analysis was performed to assess the role of prefrontal cortex in processing multiple relations simultaneously (i.e., relational integration). This analysis produced two main findings. First, activation occurred in frontopolar prefrontal cortex (BA 10), reflecting the internal generation of relations required to form complex analogies. Interestingly, this activation only occurred for 2-relational problems, not for 1-relational problems, suggesting that frontopolar cortex is important for processing complex relational structures. Second, activations also occurred in right DLPFC (BA 46), which reflected greater manipulation of externally presented information in more complex problems. Other studies using the Raven's Progressive Matrices task have found similar results, and have also reported bilateral posterior parietal activations (BA 7).

In another study, individuals received a source picture of colored geometric shapes, followed by a target picture of colored geometric shapes. Pictures that did not share similar geometric shapes but that did share the same system of abstract visuospatial relations were also presented. Individuals judged whether each source-target pairing was analogous (analogy condition) or identical (literal condition). Analogical reasoning (analogy minus literal) recruited the dorsomedial frontal cortex (BA 8) and left-hemisphere regions, including frontopolar (BA 10), inferior frontal (BA 44, BA 45, BA 46, and BA 47), and middle frontal (BA 6) cortices, and also inferior parietal cortex (BA 40). These findings suggest that analogical reasoning is mediated by a predominantly left-hemisphere frontoparietal system.

Another study systematically evaluated the component processes of analogical reasoning. Specifically, this study assessed the neural systems that underlie (1) the storage of abstract relations in working memory and (2) the process of integrating abstract relations to form analogies. These researchers also found that analogical reasoning activated a left frontoparietal system, with some regions of this circuit mediating working memory processes and others mediating abstract relational integration. In particular, left frontopolar regions (BA 9/10) were again central to the processing of relations that underlie analogical reasoning.

In summary, the observed findings for analogical reasoning are consistent with theories that advocate distributed rather than localized representations (e.g., dual-process, dual-code, embodied theories). As we saw, distributed frontal and posterior systems appear to play a wide variety of roles as representations are retrieved, stored, and integrated. Findings on analogical reasoning further support the cognitive demand and task specificity hypotheses (e.g., dual-process, dual-code, embodied theories). Different kinds of information recruited different prefrontal areas (e.g., external vs. internal information), and the harder the reasoning, the more areas recruited (e.g., 1-relation vs. 2-relation problems). The general importance of frontopolar cortex for generating relations internally suggests that this area is especially important for processing complex analogies.

Summary, Conclusions, and Future Directions

The findings discussed in this article support the following conclusions about the neural bases of reasoning. First, deductive reasoning does not appear to recruit a unitary neural system but instead engages different brain regions based on the particular reasoning task and materials employed. Second, inductive inferences drawn from familiar categorical syllogisms and conditional statements engage a left-hemisphere language and knowledge network implemented in frontal and temporal regions. Third, problem solving generally recruits prefrontal regions, including bilateral DLPFC and MVPFC. Fourth, causal reasoning does not engage a single neural system but instead recruits different systems based on the causal reasoning task and the consistency of causal evidence with background beliefs. Finally, analogical inference selectively recruits frontal and parietal regions, with frontopolar cortex becoming increasingly important as task complexity increases.

The neuroscience evidence reviewed helps evaluate current psychological theories. First, reasoning typically recruits broadly distributed neural systems, providing evidence inconsistent with the relatively localized predictions of modularity theory. Possible exceptions include the importance of certain frontal regions for problem solving (left DLPFC) and analogical reasoning (frontopolar cortex). Second, the results demonstrate that reasoning does not solely recruit left-hemisphere regions for language and rule-based operations. Instead, reasoning often engages bilateral and posterior regions beyond those implicated by amodal and mental logic theories of reasoning. Third, the predominantly right-hemisphere system predicted by the mental-models theory across both deduction and induction is inconsistent with many patterns of left-hemisphere and bilateral activation observed across studies. Fourth, dual-process, dual-code, and embodied theories are generally consistent with the reviewed findings. These theories receive support because they predict the presence of (1) broadly distributed neural systems and (2) multiple reasoning systems (e.g., associative vs. rule, language vs. knowledge). These theories also receive support from the many studies that exhibit effects of task specificity and cognitive demand. As task conditions change, so do the neural systems that represent and process the relevant information.

In general, reasoning tends to recruit broadly distributed and diverse neural systems. The difficulty of establishing specific neural systems for a given type of reasoning (e.g., deduction) strikes us as one major challenge for future research. Does a particular type of reasoning consistently activate a specific neural circuit across wide variation in tasks and materials? If so, what is this circuit? Another major challenge is that the multifaceted nature of the neurobiological evidence and cognitive theories results in a many-to-many mapping: neural regions often serve multiple cognitive functions that can be mapped onto multiple cognitive theories. Sorting out the functional roles of particular brain areas and their roles in psychological theories of reasoning should be another major goal in this research area.

Many other challenges also await future research in this area. For example, do the results from the laboratory tasks reviewed here generalize to everyday reasoning tasks? Rather than occurring in a vacuum, everyday reasoning often occurs in social situations and is associated with emotional affect. What are the neural bases of reasoning under these conditions? How do social and emotional processes modulate reasoning? Future research should also address the important distinction between reasoning to a conclusion versus recognizing a conclusion. Because many real-world situations require that people generate valid conclusions (not just recognize them), future neuroscience research should assess the neural bases of this process. Future research should also continue to address the role of task difficulty in reasoning. Does task difficulty result in the recruitment of more brain regions, or are the same regions activated more intensely? Future research should also address the effect of learning on the neural systems that underlie reasoning. Do novices and experts engage similar neural systems during reasoning?

Finally, researchers should develop psychological theories that motivate fine-grained neurobiological predictions, and should design experiments that distinguish between theories, rather than simply attempting to confirm one. Although much progress has been made in framing the issues and establishing preliminary evidence, neuroscience research on reasoning is still in its infancy. We anticipate major advances in coming years.

Acknowledgments

This work was supported by National Science Foundation Grants DGE-0536941 and DGE-0231900 to A K Barbey, and by National Science Foundation Grant BCS-0212134 and Defense Advanced Research Programs Agency Contract FA8650–05-C-7256 to L W Barsalou.

See also: Decision-Making and Vision; Executive Function and Higher-Order Cognition: Assessment in Animals; Games in Monkeys: Neurophysiology and Motor Decision-Making; Memory Representation; Prefrontal Cortex: Structure and Anatomy; Prefrontal Cortex; Referentiality and Concepts in Animal Cognition; Reward Decision-Making.

Further Reading

Barbey AK and Sloman SA (in press) Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*.

Barsalou LW (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660.

Braine MDS (1990) The 'natural logic' approach to reasoning. In: Overton WF (ed.) *Reasoning, Necessity, and Logic: Developmental Perspectives*, pp. 135–158. Mahwah, NJ: Erlbaum Press.

Christoff K and Gabrieli JDE (2000) The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology* 28: 168–186.

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

Colvin MK, Dunbar K, and Grafman J (2001) The effects of frontal lobe lesions on goal achievement in the water jug task. *Journal of Cognitive Neuroscience* 13: 1129–1147.

Flombaum JI, Santos LR, and Hauser MD (2002) Neuroecology and psychological modularity. *Trends in Cognitive Sciences* 6: 106–108.

Fonlupt P (2003) Perception and judgment of physical causality involve different brain structures. *Cognitive Brain Research* 17: 248–254.

Fugelsang J and Dunbar K (2004) A cognitive neuroscience framework for understanding causal reasoning and the law. *Philosophical Transactions of the Royal Society of London. Series B* 359: 1749–1754.

Goel V (2002) Planning: Neural and psychological. In: Nadel L (ed.) *Encyclopedia of Cognitive Science,* vol. 3, pp. 697–703. New York: Macmillan.

Goel V (2005) Cognitive neuroscience of deductive reasoning. In: Holyoak K and Morrison R (eds.) *Cambridge Handbook of Thinking and Reasoning*, pp. 475–492. Cambridge University Press: New York.

Goel V, Buchel C, Frith C, et al. (2000) Dissociation of mechanisms underlying syllogistic reasoning. *NeuroImage* 12: 504–514.

Goel V and Dolan RJ (2003) Explaining modulation of reasoning by belief. *Cognition* 87: B11–B22.

Green AE, Fugelsang JA, Kraemer DJM, et al. (2006) Frontopolar cortex mediates abstract integration in analogy. *Brain Research* 1096: 125–137.

Holyoak KJ (2005) Analogy. In: Holyoak KJ and Morrison RG (eds.) *Cambridge Handbook of Thinking and Reasoning*, pp. 117–142. Cambridge, UK: Cambridge University Press.

Johnson-Laird PN (1983) *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge, UK: Cambridge University Press.

Novick LR and Bassok M (2005) Problem solving. In: Holyoak KJ and Morrison RG (eds.) *Cambridge Handbook of Thinking and Reasoning*, pp. 321–349. New York: Cambridge University Press.

Osherson D, Perani D, Cappa S, et al. (1998) Distinct brain loci in deductive versus probabilistic reasoning. *Neuropsychologia* 36: 369–376.

Over DE (2003) From massive modularity to meta-representation: The evolution of higher cognition. In: Over DE (ed.) *Evolution and the Psychology of Thinking: The Debate*, pp. 121–144. New York: Psychology Press.

Paivio A (1986) *Mental Representations*. New York: Oxford University Press.

Parsons LM and Osherson DN (2001) New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cerebral Cortex* 11: 954–965.

Roger ME, Fugelsang JA, Dunbar KN, et al. (2005) Dissociating processes supporting causal perception and causal inference in the brain. *Neuropsychology* 19: 591–602.

Wharton CM, Grafman J, Flitman SS, et al. (2000) Toward neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology* 40: 173–197.

Whitaker H, Savary F, Markovits H, et al. (1991) Inference deficits after brain damage. Paper presented at the International Neuropsychological Meeting, San Antonio, 1991.