

69 Prefrontal Substrate of Human Relational Reasoning

BARBARA J. KNOWLTON AND KEITH J. HOLYOAK

ABSTRACT Relational reasoning, including the distinctively human capacity to see analogies between disparate situations, requires the ability to mentally represent and manipulate the relationships among concepts. Over the past decade, studies of cognitive development, aging, and neurological disease have supported the hypothesis that the prefrontal cortex plays a critical role in relational reasoning. Analysis of the component processes of relational reasoning has motivated neuroimaging studies linking these components to distinct neural substrates. Studies using diverse reasoning tasks have converged on the conclusion that frontopolar cortex responds specifically when multiple relations must be integrated to solve a problem. Other subregions in the inferior and middle frontal gyri appear to be critical for resolving interference from distracting elements of the problem and for control of working memory. Mapping the components of relational reasoning is an essential first step toward understanding how the architecture of the prefrontal cortex supports human thinking.

Humans, more than any other species, are able to cope with novel problems that arise across a wide range of domains. This capacity depends in part on *role-based relational reasoning*—the ability to draw inferences about entities based on the roles they fill in relations, where the roles are not predictable by features of the entities and the relations cannot be reduced to roleless chunks (Penn, Holyoak, & Povinelli, 2008). A canonical example of role-based relational reasoning is reasoning by analogy, which enables detection of higher-order similarities between superficially dissimilar situations (Gick & Holyoak, 1980; Gentner, 1983; Holyoak, 2005). For example, modern medicine originated with Pasteur's development of the germ theory of disease by analogy to the known role of microorganisms in fermentation of grapes, coupled with Lister's extension of Pasteur's analogy to explain infections (Thagard, 1996). Analogy requires representing and integrating relational knowledge in working memory, while coping with interference from salient distracting information. For example, Pasteur grasped that microorganisms need to *be alive* and *in contact with* grapes in order to *cause* fermentation; by integrating these relations and mapping them onto what he knew about certain silk-

worm diseases, he was able to infer that germs might be their cause (and hence the diseases might be prevented by killing the germs or blocking their contact with silkworms). At the same time, he had to ignore many salient but irrelevant dissimilarities between the two analogs (e.g., silkworms are animals, grapes are not).

While prefrontal cortex (PFC) has long been associated with problem solving and fluid intelligence, in the past decade there has been a surge in work specifically relating the PFC to processes supporting relational reasoning. In 1995, Robin and Holyoak advanced the thesis that the PFC is responsible for the creation and maintenance of explicit relational representations that guide thought and action. At that time, however, few or no neuropsychological or neuroimaging data were available to directly connect relational processing with the human PFC. Moreover, although several major computational models of analogical reasoning had been developed in cognitive science (e.g., Falkenhainer, Forbus, & Gentner, 1989; Holyoak & Thagard, 1989), none even attempted to incorporate constraints based on what was known about working memory, interference control, and their neural substrates.

Today the picture is very different. Behavioral studies with children (Richland, Morrison, & Holyoak, 2006), young adults (Cho, Holyoak, & Cannon, 2007), and older adults (Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004; Viscontas, Holyoak, & Knowlton, 2005) have teased apart some of the major component processes in relational reasoning. Computational models of analogy have incorporated working-memory constraints (Halford, Wilson, & Phillips, 1998; Hummel & Holyoak, 1997, 2003). It has been shown that "virtual brain damage" in a neural-network model can simulate changes in reasoning related to normal aging (Viskontas et al., 2004), as well as reasoning deficits in patients with damage to their frontal or temporal cortices (Morrison et al., 2004). In addition to neuropsychological studies of relational reasoning conducted in our lab (Krawczyk et al., 2008; Morrison et al., 2004; Waltz et al., 1999, 2004), neuroimaging studies from multiple labs have identified subregions of PFC associated with integrating multiple relations (Bunge, Wendelken, Badre, & Wagner, 2005; Christoff et al., 2001; Green, Fugelsang, Kraemer, Shamosh,

BARBARA J. KNOWLTON and KEITH J. HOLYOAK Department of Psychology, University of California, Los Angeles, California

& Dunbar, 2006; Kroger et al., 2002; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997; Wendelken, Nakhbenko, Donohue, Carter, & Bunge, 2008), as well as with interference control (Cho et al., 2009).

In this chapter we discuss how a clearer picture of the neural basis of relational reasoning has emerged from current conceptions of its component processes.

Component processes of relational reasoning

RELATIONAL INTEGRATION AND INTERFERENCE CONTROL As Robin and Holyoak (1995) argued, relational reasoning appears to require processes closely related to those assoc-

iated with the operation of the PFC. As the name implies, role-based relational reasoning requires drawing inferences about entities based on the *roles* they fill, rather than on direct similarity. A specific task that instantiates these requirements is illustrated in figure 69.1, which depicts examples of “scene-analogy” problems developed by Richland and colleagues (2006) for use with children as young as 3–4 years. For each pair of pictures, children were asked to identify the object in the bottom picture that “goes with” the object indicated by an arrow in the top picture. In some problems, such as that shown in figure 69.1*B*, the child is confronted with a conflict between two possible answers, one relational and one based on perceptual and/or semantic

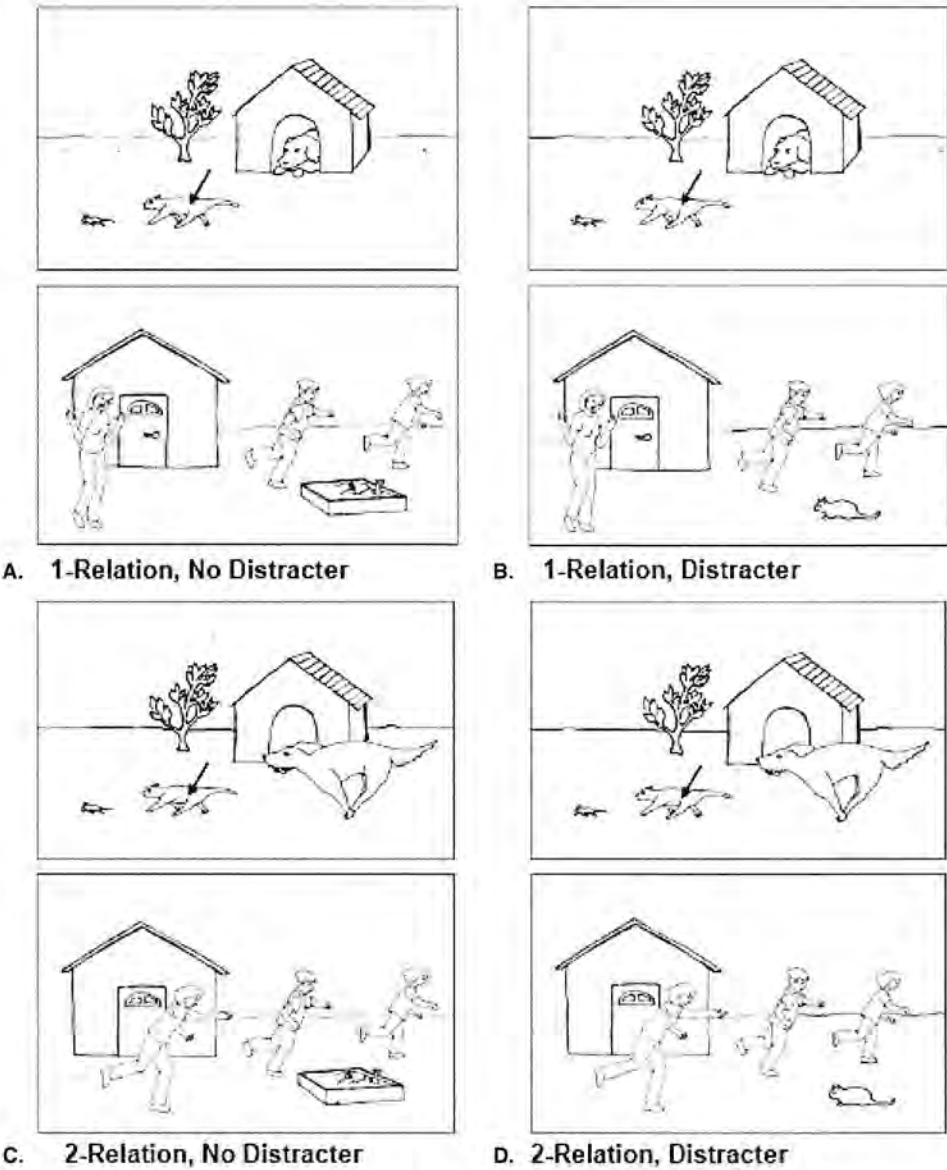


FIGURE 69.1 Example set of scene pairs constructed by Richland, Morrison, and Holyoak (2006). (A) 1-relation/no distracter; (B) 1-relation/distracter; (C) 2-relation/no distracter; (D) 2-relation/distracter. (Reprinted with permission from Richland, Morrison, & Holyoak, 2006.)

similarity. The cat in the top picture perceptually resembles the cat in the bottom picture, but plays a role (chasing a mouse) that parallels the role played by the boy in the bottom picture (chasing a girl). Richland and colleagues found that young children were less likely to give the relational response when an alternative based on direct similarity was available (figure 69.1*B*) than when no such distracter object was present (figure 69.1*A*, where in the bottom picture the cat has been replaced by a sandbox).

Relational reasoning varies in its complexity, which has been linked to the number of relational roles relevant to an inference (Halford, 1993; Halford et al., 1998). A critical distinction is whether a single relation is sufficient to determine the role-based inference (figure 69.1*A,B*), or whether multiple relations must be integrated to derive an inference (figure 69.1*C,D*). For example, in figure 69.1*D* it is not sufficient simply to relate the cat to another “chaser” in the bottom picture, as in the latter both the boy and the woman are chasing someone. Rather, the cat corresponds specifically to the boy because each is both being chased (by the dog and the woman, respectively) as well as chasing (the mouse and the girl, respectively). Richland and colleagues (2006) found that preschool children gave fewer relational responses when either a similar distracter was present in the bottom picture (figure 69.1*B,D* versus *A,C*) or when two relations had to be integrated (figure 69.1*C,D* versus *A,B*). By age 13–14 years—roughly the age at which the PFC has undergone substantial further maturation (Giedd, 2004)—children reliably gave the relational response even when multiple relations had to be integrated and a similar distracter was present.

LISA: A NEUROCOMPUTATIONAL MODEL OF RELATIONAL REASONING Research on the development of analogical

reasoning thus suggests that relational reasoning depends on two major component processes: integration of multiple relations and the capacity to cope with interference from salient but relationally irrelevant information. Most computational analogies of analogy have not considered how analogical reasoning might be implemented in the brain, and traditional neural-network models of cognition encounter severe difficulty in representing explicit relations and the binding of objects into relational roles (Doumas & Hummel, 2005). However, a neural-network model of the component processes underlying relational reasoning has been developed in recent years. LISA (*Learning and Inference with Schemas and Analogies*; Hummel & Holyoak, 1997, 2003) uses a representation of knowledge based on a hierarchy of distributed and localist units (see figure 69.2). LISA’s representational scheme codes relations as sets of roles (e.g., for “cat chases mouse,” the cat fills the role of “chaser” and the mouse the role of “chased”). The model uses synchrony of firing to bind distributed representations of relational roles (e.g., chaser) to distributed representations of their fillers (e.g., cat). The process of “thinking about” a proposition entails keeping separate role-filler bindings firing out of synchrony with one another. In LISA, working memory is necessarily capacity-limited: It is only possible to keep a finite number of role-filler bindings simultaneously active and out of synchrony with one another (for details see appendix A in Hummel & Holyoak, 2003).

LISA represents propositions using a hierarchy of distributed and localist units. Figure 69.2*B* provides a schematic representation of LISA’s architecture as applied to one of the scene-analogy problems used by Richland and colleagues (2006). At the bottom of the hierarchy, *semantic* units (small circles in figure 69.2*B*) represent objects and relational roles in a distributed fashion. For example, consider the

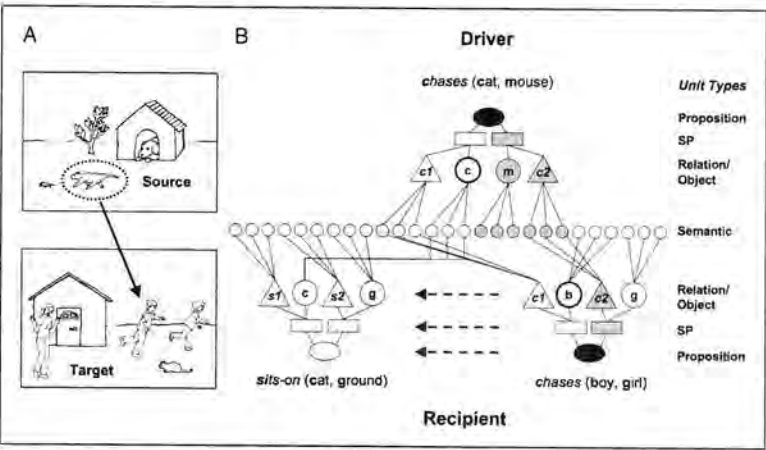


FIGURE 69.2 (A) Example of 1-relation/distracter scene-analogy problem (Richland et al., 2006). (B) LISA architecture as applied to this problem. In order for a reasoner to select the boy in the target as the correct analogical mapping to the cat in the source,

units in the recipient representing *chases* (boy, girl) must inhibit corresponding units in the propositional structure containing the featurally similar “sitting cat” distracter. (Reprinted with permission from Morrison, Doumas, & Richland, 2006.)

proposition *chase* (cat, mouse). Each role of the *chase* relation would be represented by units coding for its semantic content (e.g., *aggressor* for the first role, *victim* for the second, and *pursuit* for both). Similarly, the objects “cat” and “mouse” would be represented by units specifying their meaning (e.g., cat: *animal*, *pet*, *soft*). *Predicate* and *object* units (triangles and large circles, respectively, in figure 69.2B) represent relational roles and their object fillers, and have bidirectional excitatory connections to the corresponding semantic units. Subproposition (SP) units (rectangles in figure 69.2B) bind roles to their arguments, and have bidirectional connections to the corresponding predicate and object units. In the case of *chase* (cat, mouse), one SP would bind “cat” to the first role of *chase*, and another would bind “mouse” to the second. At the top of the hierarchy, proposition (P) units bind role-filler bindings into complete propositions by way of excitatory connections to the corresponding SPs. A complete analog (i.e., situation, story, or scene) is represented by the collection of semantic, predicate, object, SP, and P units that collectively code the propositions in that analog. The semantic units permit the units in one analog to communicate with the units in others.

To generate an analogical mapping, units representing one analog (the *driver*) are activated in working memory, and reasoning proceeds by passing activation from these units through distributed semantic units to units representing the *recipient* analog in long-term memory. As units in the recipient analog are fired, they enter working memory. LISA postulates a set of *mapping connections* between units of the same type in separate analogs. These connections grow whenever the corresponding units are active simultaneously and thereby permit LISA to rapidly learn the correspondences between structures in separate analogs.

The basic processes of LISA are closely related to the functions of the PFC. Hummel and Holyoak (1997, 2003) hypothesized that the rapid learning of mapping connections, which is critical to relational integration, is an important function of working memory as implemented in prefrontal cortex (cf. Assad, Rainer, & Miller, 1998). Inhibitory control, which is also considered an important function of prefrontal cortex (Miller & Cohen, 2001; Shimamura, 2000), plays a central role in several aspects of LISA. These include (1) LISA’s ability to select items for placement into working memory, (2) its working memory capacity for role-filler bindings, (3) its ability to control the spreading of activation in the recipient (i.e., its ability to disambiguate which elements of the recipient correspond to the active units in the driver), (4) its ability to use competition among mapping connections to enforce structural constraints on the discovery of analogical mappings, particularly the constraint that mappings tend to be one-to-one, and (5) its ability to select a relation-based response despite the availability of a salient but superficial distracter.

Importantly, LISA predicts that relational integration and inhibitory control are distinct but closely linked processes. Relational integration requires the ability to rapidly learn mapping connections, but inhibitory control is essential to set up the conditions for successful learning (i.e., simultaneous activation of objects that fill parallel relational roles). As we will see, LISA’s conception of the relationship between relational integration and inhibitory control is consistent with recent functional imaging data that indicate how these processes map onto subregions of PFC.

Neuropsychological evidence for the role of PFC in relational reasoning

One of the major sources of evidence concerning the dependence of relational reasoning on the integrity of the PFC comes from neuropsychological studies of patients diagnosed with frontotemporal lobar degeneration (FTLD), a dementia subtype distinct from Alzheimer’s disease (Brun, 1993; Snowden, Neary, & Mann, 2007). FTLD occurs in three variants: a progressive aphasia, in which areas involved in language production are primarily affected initially; a frontal variant, which is characterized in early stages by prefrontal atrophy; and a temporal variant (also known as *semantic dementia*) in which degeneration includes left anterior temporal cortex (Hodges, Patterson, Oxbury, & Funnell, 1996). Frontal-variant FTLD provides a model for investigating what reasoning processes are dependent on PFC, while the temporal variant provides a closely matched control with the same underlying disease.

RELATIONAL COMPLEXITY AND THE PFC Waltz and colleagues (1999) examined performance of FTLD patients and age-matched normal control subjects on simple reasoning tasks, using closely matched variants of problems that differed specifically in whether or not success required integration of multiple relations. They hypothesized that patients with prefrontal cortical dysfunction would exhibit impaired performance when asked to integrate multiple relations, yet would perform normally when only one relation needed to be considered. The performance of frontal-variant FTLD patients was compared to that of temporal-variant patients, as well as age-matched healthy controls.

Figure 69.3 illustrates how relational complexity was manipulated for problems adapted from the Raven’s Standard Progressive Matrices Test (RPM), which has long been used as a measure of reasoning ability (Raven, 1941). Non-relational problems (level-0 complexity) involved a visual pattern, with a blank space in the bottom right-hand corner (see figure 69.3.4). On these problems, the participant could simply pattern-match to select the correct completion. Each one-relation problem (level-1 complexity) involved a 2×2 matrix that required processing one relational change over

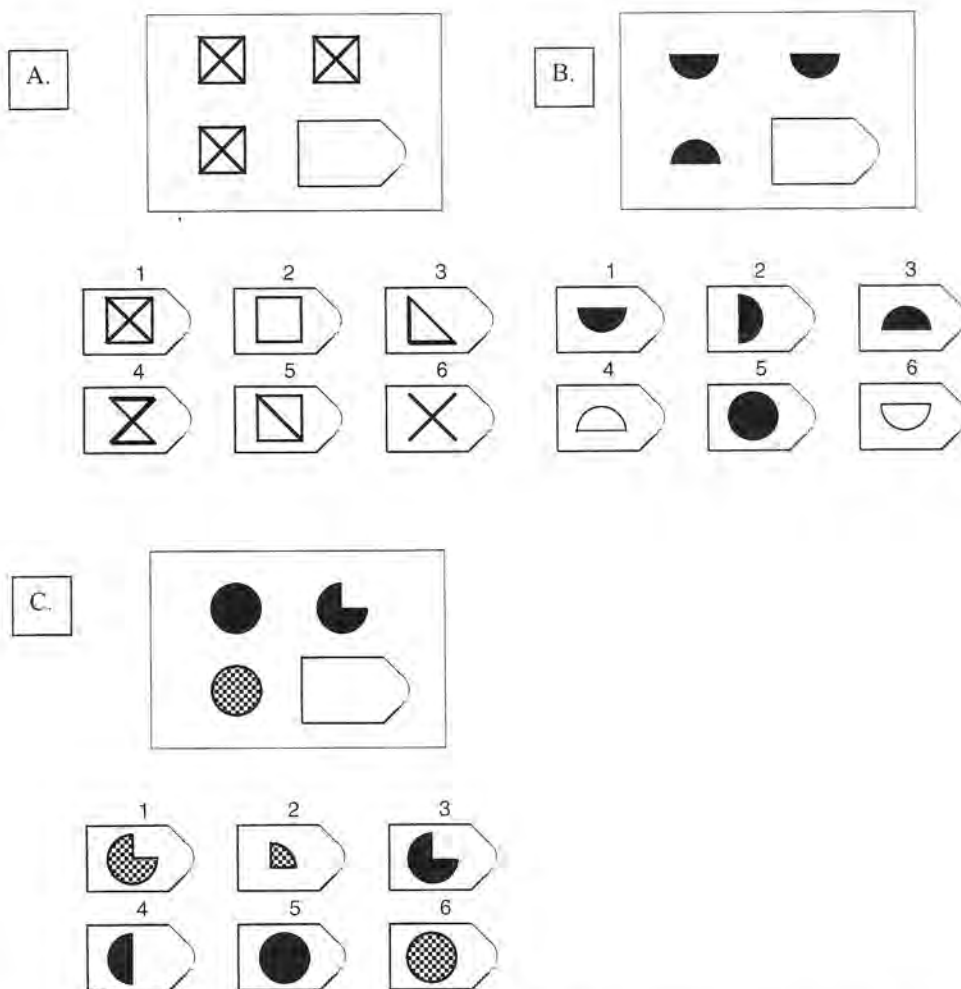


FIGURE 69.3 Examples of problems adapted from the Raven Standard Progressive Matrices Test by Waltz and colleagues (1999). (A) Nonrelational problem (level 0), requiring only perceptual matching (correct response is choice 1). (B) One-relation problem (level 1), requiring processing of the transformation along the vertical dimension only (reflection across the x -axis) in order to choose

the correct alternative (choice 3). (C) Two-relation problem (level 2), requiring integration of the relation along the vertical dimension (solid to checked pattern) and the relation across the horizontal dimension (removal of the upper-right quadrant) in order to choose the correct response (choice 1). (Reprinted with permission from Waltz et al., 1999.)

either the horizontal or vertical dimension; the other dimension was constant (figure 69.3B). Two-relation problems (level-2 complexity) required integrating two relational changes over the horizontal and vertical dimensions, respectively (figure 69.3C). Thus, although the basic form of the task was constant across the three types of matrix problems, only the two-relation problems necessitated relational integration.

Figure 69.4 presents the results obtained by Waltz and colleagues (1999) for the matrix problems. The temporal-variant FTLN patients and the normal controls achieved very high accuracy at all levels of relational complexity. The frontal-variant patients also performed at a high level for level-0 and level-1 problems; however, their performance plunged dramatically for level-2 items (just 11% correct, not different from chance). At the same time, the performance

of frontal-variant patients was superior to that of temporal-variant patients on a test of recognition memory and on tests dependent on semantic knowledge. The resulting double dissociation between relational reasoning versus both episodic memory and semantic knowledge rules out a general “difficulty” factor as the source of the prefrontal group’s impairment.

The frontal-variant patients also were completely unable to solve transitive-inference problems that required integration of two relations. Reasoning deficits of a similar but less pronounced nature have also been observed in patients with Alzheimer’s disease with pronounced frontal signs (Waltz et al., 2004). Taken together, these findings indicate that the human PFC plays an essential role in relational reasoning—specifically, in the integration of multiple relations. Furthermore, the role of the PFC in relational integration

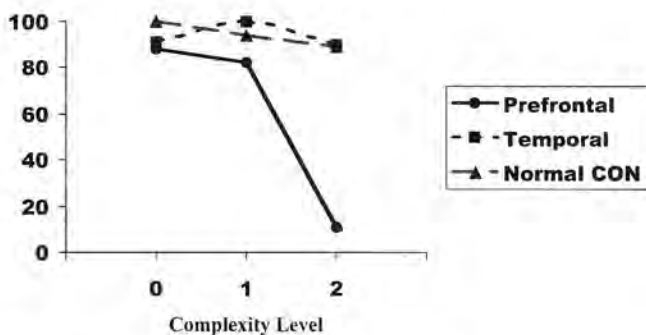


FIGURE 69.4 Accuracy on matrices test (Waltz et al., 1999). Groups show similar performance in the solution of nonrelational problems (level 0), as well as those requiring maintenance of a single relation (level 1); but patients with prefrontal damage show catastrophic impairment in the ability to solve problems requiring the integration of multiple relational premises (level 2). (Reprinted with permission from Waltz et al., 1999.)

was shown to generalize across both inductive (RPM) and deductive (transitive inference) reasoning tasks.

INTERFERENCE CONTROL AND THE PFC Other recent neuropsychological studies using FTLN patients have examined the role of the PFC in controlling interference from distracting information during analogical reasoning. Morrison and colleagues (2004) tested both frontal- and temporal-variant FTLN, as well as age-matched controls, on a verbal analogy task. Four-term analogy problems of the form A:B::C:D or D' were employed, where D is the analogical answer and D' is a nonanalogical foil (adapted from Sternberg & Nigro, 1980). A semantic facilitation index (SFI) was calculated for each problem to characterize the association of the correct relational pair (C:D) relative to the distracter pair (C:D'). For example, for the problem PLAY:GAME::GIVE: ? (1) PARTY (2) TAKE, the C:D pair (GIVE:PARTY, the correct analogical answer) is less associated than is the C:D' pair (GIVE:TAKE, the non-analogical foil), yielding a negative SFI for the problem. The problems were divided into those with negative SFI, neutral SFI, and positive SFI in order to examine the effect of semantic interference on the ability to identify the analogical answer.

Morrison and colleagues (2004) predicted that because these 4-term analogy problems were based on a single, fairly simple relation between the A and B terms, frontal-variant FTLN patients should be able to perform the basic analogical mapping needed despite their diminished working memory. However, if the PFC is also critical for interference resolution, then these patients should be selectively impaired on problems in which the D' distracter is a strong competitor to the analogical choice, D. In the positive and neutral SFI conditions, the analogical answer (D) does not face competition from an alternative (D') that is more strongly associated

with the C term. Accordingly, the analogical answer can simply be activated and produced as a response. However, in the negative SFI condition the D' foil is in fact more strongly associated with C than is the analogical response D. It follows that in order to make the analogical response for these problems, it will be necessary not just to activate the D response, but also to inhibit the semantically related D' response.

Accordingly, because of their postulated deficits in inhibitory control, it was predicted that frontal-variant FTLN patients would be selectively impaired in the negative SFI condition relative to the positive and neutral SFI conditions. In contrast, it was predicted that temporal-variant FTLN patients would show a more uniform decline in verbal analogy performance across all three conditions because of their loss of the conceptual information necessary to encode the relations in the analogy problem. Both of these patterns were in fact observed by Morrison and colleagues (2004). A more recent study using 4-term picture analogies also found that frontal-variant FTLN patients are especially impaired on problems that include semantically related distracters (Krawczyk et al., 2008).

The deficits in the frontal- and temporal-lobe patient groups that Morrison and colleagues (2004) found with the verbal analogy task were modeled using LISA. It proved possible to simulate the observed pattern of frontal-lobe deficits by impairing the rate of rapid learning of analogical connections in LISA's working memory, coupled with reduction of a parameter for inhibitory control. Both the rapid learning of new connections in working memory and inhibitory control appear to be key functions of prefrontal cortex (Miller & Cohen, 2001; Shimamura, 2000). When both these functions (not either one alone) were impaired in LISA, the model yielded the selective impairment on negative SFI problems shown by frontal-lobe patients. Thus LISA would predict that distinct prefrontal regions would be activated during analogical reasoning corresponding to these two components. When the extent of semantic death (loss of connections between semantic units representing a relational role and the predicate unit for that role) was increased in LISA, thereby modeling loss of conceptual knowledge in anterior temporal cortex, the simulation yielded the pattern of impairment found for temporal-variant FTLN patients: a relatively uniform decrease in accuracy across all verbal analogy problems, regardless of SFI condition.

RELATIONAL REASONING IN THE AGING BRAIN The PFC has been shown to be vulnerable to the effects of aging (Raz et al., 1997). Decreased PFC function in older adults may affect the ability to cope with relational complexity and interference during reasoning. Viskontas and colleagues (2004) manipulated both factors using a set of analogy problems based on cartoons of human figures, each defined

by four binary-valued characteristics (height, girth, clothing color, and gender). Figure 69.5 illustrates examples of these People Pieces Analogies (PPA; Sternberg, 1977; Morrison, Holyoak, & Truong, 2001; Cho et al., 2007). In the PPA task, subjects are asked to verify as quickly as possible whether the relationship between the A:B pair corresponds to the relationship between the C:D pair with respect to selected dimension(s). This task has a major advantage over other paradigms because visual complexity is controlled by having all problems, regardless of the level of relational

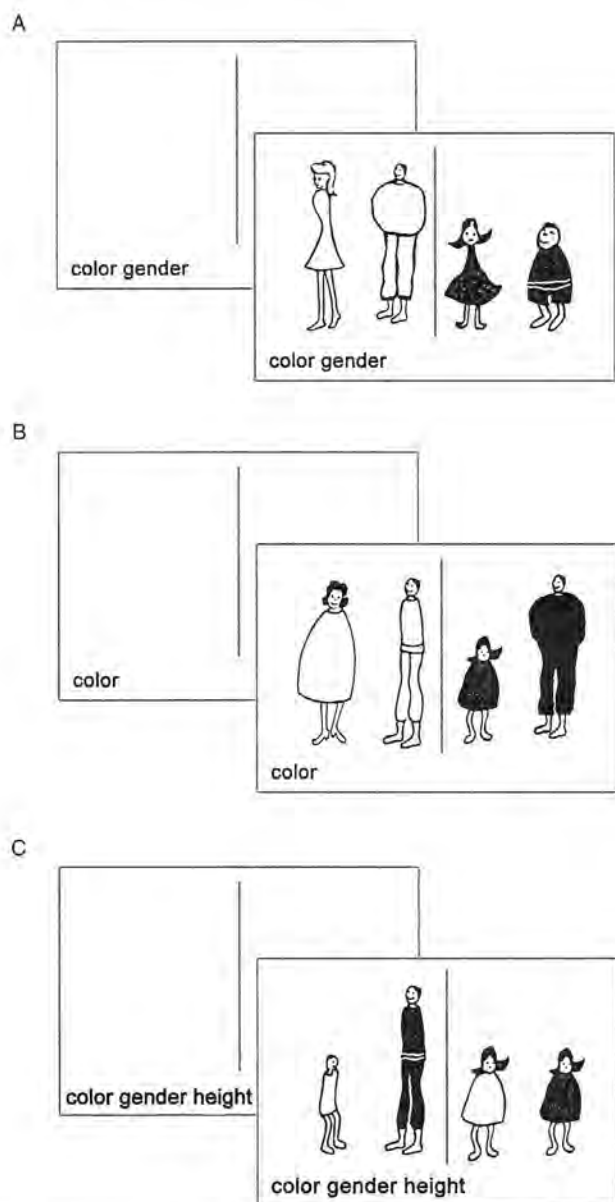


FIGURE 69.5 Example problems from People Pieces Analogy task (Viskontas et al., 2004). (A) two relations, none to inhibit; (B) one relation to attend to, two to inhibit; (C) problem for which correct response is “false.” (Reprinted with permission from Vis Kontas, Morrison, Holyoak, Hummel, & Knowlton, 2004.)

complexity, consist of four characters from the basic set of 16. Before each trial in the PPA task participants see a list of the traits that they are to attend to on that trial. They then see an analogy problem that they have to judge as true or false based on the dimensions specified as relevant. The number of traits to which participants attend can be varied from one to four, thus manipulating the relational complexity of the problem. In addition, 1- and 2-relation problems may have 0, 1, or 2 irrelevant dimensions that need to be inhibited to avoid interference with the analogical answer. “False” trials are constructed by introducing the wrong value on one relevant dimension.

Viskontas and colleagues (2004) administered the PPA tasks to three groups of adult subjects: younger, middle-aged, and older (mean ages of approximately 20, 50, and 75 years, respectively). The mean response times (RTs) for correct “true” trials are shown in figure 69.6 (left). For both young and older subjects, RTs were significantly increased for greater levels of relational complexity (compare figure 69.6A to 69.6C). For young subjects RTs increased only slightly when distracting information was present. In older adults, however, distracting information had a profound effect on response time when it was necessary to integrate more than one relation (figure 69.6C). Age-related impairments in relational integration have also been obtained with other deductive and inductive reasoning tasks (Viskontas et al., 2005).

Viskontas and colleagues (2004) showed that the LISA model could successfully capture the interaction among age, complexity, and the amount of distracting information simply by varying an inhibition parameter. Reducing the model’s inhibition parameter has the effect of increasing sensitivity to distracting information, particularly when multiple relations need to be mapped in working memory (see figure 69.6, right). Such loss of inhibitory control for older adults is consistent with the decline in prefrontal functions that accompanies cognitive aging. In younger adults, performance on analogy problems that require interference control can be impaired by imposing dual-task conditions (Morrison et al., 2001; Waltz, Lau, Grewal, & Holyoak, 2000), by delaying the cue signaling the relevant dimensions on a trial (Cho et al., 2007), or by inducing anxiety prior to administering the analogy task (Tohill & Holyoak, 2000). All these factors plausibly act by imposing excessive load on working-memory processes dependent on the PFC.

Functional imaging of component processes in relational reasoning

Neuropsychological evidence has established the critical role of the PFC in relational reasoning, and behavioral and computational modeling studies point to the distinctive roles of relational integration and inhibitory control. In recent years, considerable progress has been made in localizing specific

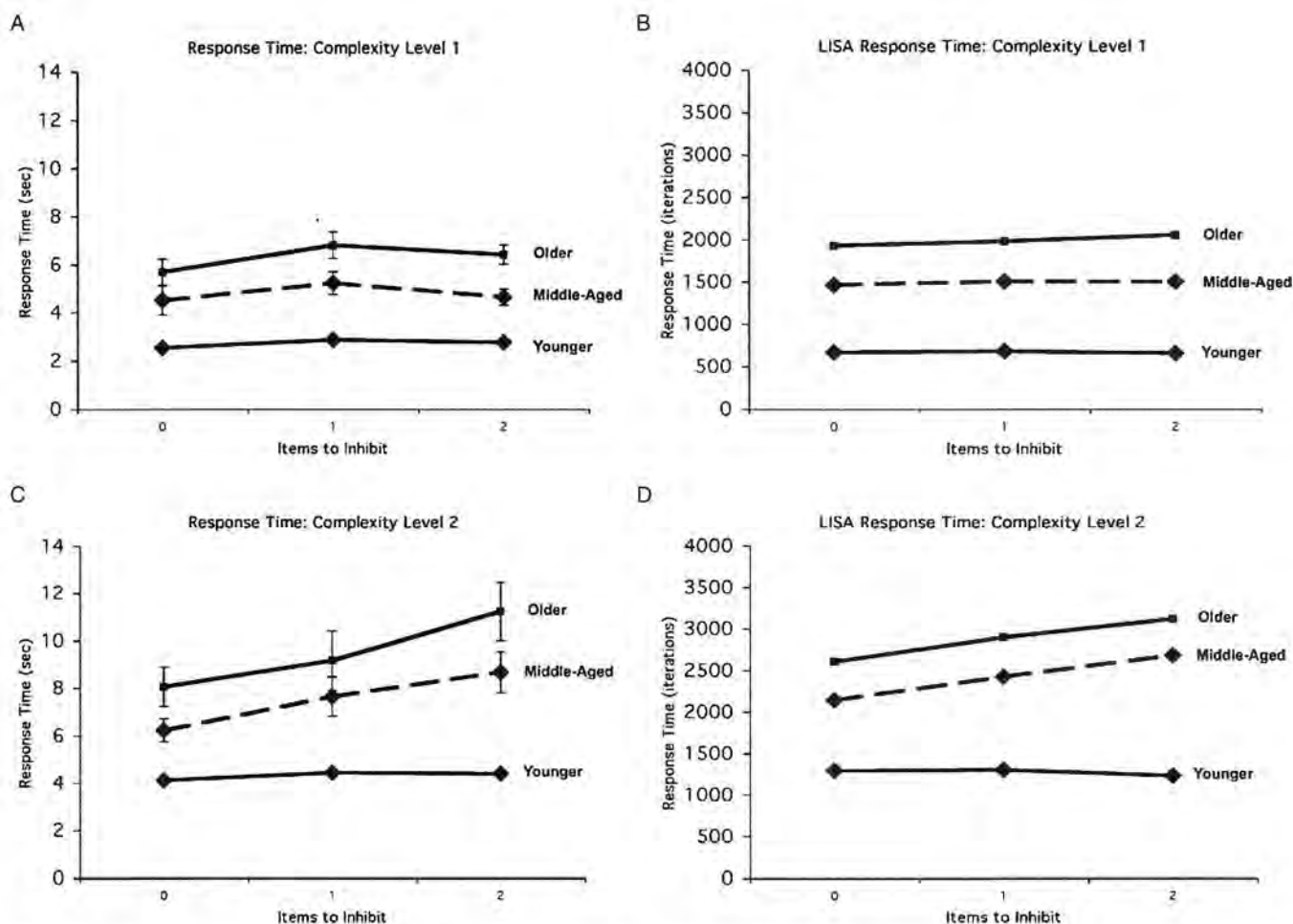


FIGURE 69.6 (A) Response time in the People Pieces Analogy task for 3 levels of inhibition at complexity level 1 for younger, middle-aged, and older groups; (B) corresponding LISA simulations; (C)

human data for complexity level 2; and (D) corresponding LISA simulations. Error bars depict standard error of the mean. (Reprinted with permission from Viskontas et al., 2004.)

PFC and other brain areas associated with relational reasoning using PET (Wharton et al., 2000) and fMRI.

RELATIONAL INTEGRATION Several fMRI studies have manipulated relational complexity using variants of RPM problems, similar to those used by Waltz and colleagues (1999) with FTLD patients. With healthy young adults, more complex problems can be used (typically 3×3 matrices with up to four dimensions of variation; Kroger et al., 2002). For matrix problems, relational integration has been shown to consistently activate prefrontal regions. In particular, bilateral middle (MFG) and inferior (IFG) frontal gyri, as well as parietal and occipital regions, have been found to increase activity when multiple relations must be integrated in order to arrive at a solution, compared to problems that require processing of only a single relation (Prabhakaran et al., 1997; Christoff et al., 2001; Kroger et al., 2002).

Among these regions, which constitute a network commonly activated in visuospatial working memory tasks, the

activation pattern of the most anterior part of the PFC has been particularly noteworthy. Christoff and colleagues (2001) found that the left frontopolar region remained preferentially activated even after controlling for the influence of increased problem-solving time. Kroger and colleagues (2002) confirmed and extended these results, providing evidence that although the left anterior prefrontal region becomes increasingly activated as more relations need to be integrated, activation in this subarea is not affected by increases in perceptual difficulty (achieved by adding visuospatial distracters to RPM problems while holding relational complexity constant). Thus the frontopolar region seems to be uniquely associated with the specific requirement of integrating multiple relations, and not general cognitive difficulty.

Similarly, studies of verbal analogical reasoning have distinguished neural substrates of reasoning from semantic processing demands within working memory. Activation in the left frontopolar region increases selectively when making

judgments of analogical similarity compared to processing of semantic associations or categories (Bunge et al., 2005; Green et al., 2006; Wendelken et al., 2008). Thus, based on a substantial body of findings involving solution of different types of relational reasoning problems, the frontopolar region seems to play a special role in the process of integrating multiple relational representations to arrive at a solution.

SEPARATING RELATIONAL COMPLEXITY AND INTERFERENCE CONTROL While several studies have examined the neural correlates of relational integration, less is known about the neural basis of interference control in the context of analogical reasoning. Interference resolution (often linked to selection and inhibitory control) has been extensively studied using a wide variety of paradigms. Studies have identified the lateral PFC, including regions in the dorsolateral PFC and inferior frontal gyrus, as important for interference resolution across diverse tasks including inhibition of a motor response, proactive interference resolution in working memory, selection among competing alternatives, and controlled semantic retrieval (Aron, Robbins, & Poldrack, 2004; Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; D'Esposito, Postle, Jonides, & Smith, 1999; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Nee, Wager, & Jonides, 2007; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Thompson-Schill et al., 2002; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001; Zhang, Feng, Fox, Gao, & Tan, 2004). It seems plausible that these regions also support interference resolution in relational reasoning.

Recently, work in our lab (Cho et al., 2009) has jointly examined the neural substrate for relational integration and for interference resolution using the PPA task described earlier (Viskontas et al., 2004). In this task, it is possible to vary both factors simultaneously while holding visuospatial complexity constant. As in previous studies using the PPA task, the subject had to determine whether the analogy between the two pairs was valid, based on a subset of trait(s) randomly selected for each trial. Participants were instructed to solve each problem based on relevant ("to-be-attended") trait(s) only, to ignore irrelevant ("to-be-ignored") traits, and to decide as quickly and accurately as possible. A trait list consisting of four words naming each trait was displayed in the center of the screen, between the A:B and C:D pairs of cartoon characters. On each 8-second trial, the target pair (A:B) appeared on the screen for 1.7 s (target phase). The trait names were all shown in black font during the target phase. After the target pair disappeared, the to-be-attended trait cue(s) turned red (cue phase) and remained on the screen. After 0.3 s, the probe pair (C:D) appeared on the right side of the screen (probe phase), and subjects were allowed a maximum of 6 s to respond with a key press. The delayed cuing of to-be-attended traits ensured that subjects

had to actively pay attention to all visual information about the A:B pair and that potential sources of interference would therefore be encoded into working memory.

Relational complexity level (number of to-be-attended traits, 1 or 3) and need for interference resolution (number of to-be-ignored traits that supported an incorrect response, 0 or 1) were manipulated in a factorial event-related fMRI design. There was a large increase in response time (RT) for PPA problems at the higher level of relational complexity. There was also an interactive effect of interference, which resulted in a reliable increase in RT only at the higher level of relational complexity. This overadditive interaction in the RT data resembled that found in previous studies using the PPA task (Cho et al., 2007; Viscontas et al., 2004). Both higher complexity and presence of interference significantly reduced accuracy in solving the analogy problems.

The fMRI analyses revealed cortical regions sensitive to increase in demands on relational integration, interference resolution, or both component processes of reasoning (see figure 69.7). In a whole-brain analysis, clusters activated by an increase in relational complexity were identified in bilateral frontal pole, as well as other regions including the MFG and IFG. In an a priori defined anatomical region-of-interest analysis of the lateral PFC comprising bilateral MFG and IFG, regions sensitive to increase in demands on interference resolution were found in bilateral MFG and IFG pars opercularis and the IFG pars triangularis in the right hemisphere. Regions sensitive to both component processes of analogical reasoning were found in bilateral MFG and IFG pars opercularis and the right IFG pars triangularis. It is possible that the neural basis for the behavioral RT interaction that was observed between relational complexity and interference may lie in this region of spatial overlap of activation. As suggested by the LISA model, inhibitory control is particularly critical when multiple relations are mapped in working memory. These overlapping regions may thus be involved in inhibitory control in working memory.

Consistent with proposals that the frontopolar region plays a specific role in reasoning when the outcomes of two or more relational comparisons must be integrated to arrive at a solution (Waltz et al., 1999; Christoff et al., 2001; Kroger et al., 2002), Cho and colleagues (2009) found signal increases in bilateral frontal pole related to increases in relational complexity but not to increased difficulty created by adding interference. By eliminating alternative explanations based on visual complexity or general cognitive difficulty, these results provide strong evidence supporting the hypothesis that the frontal pole plays a distinct role in reasoning tasks that require the integration of multiple relations.

The region-of-interest analyses found that areas of the lateral PFC that have been identified as important in executive control in a variety of cognitive tasks (dorsolateral and inferior frontal region) are also activated by the need

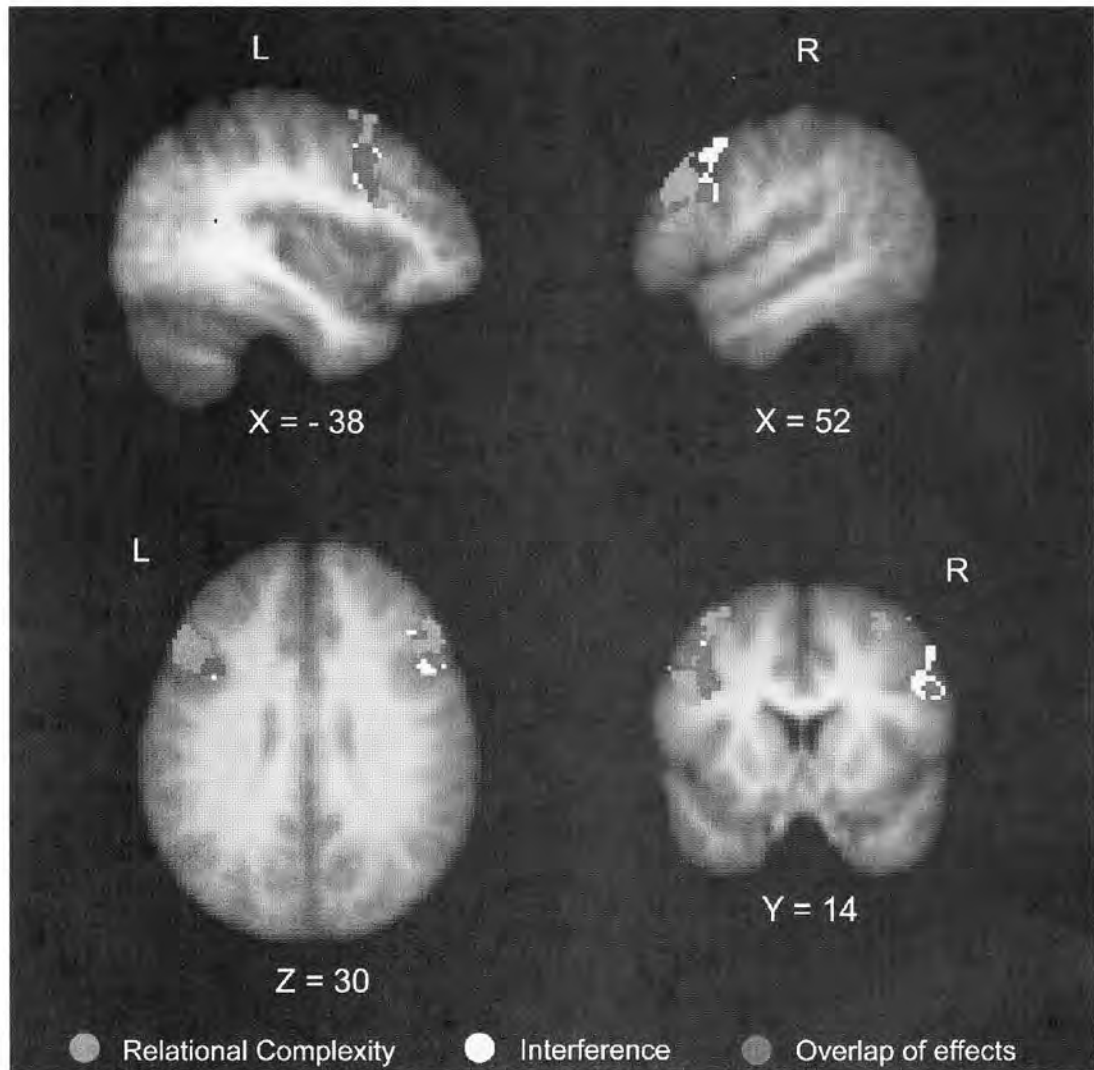


FIGURE 69.7 Neuroimaging results from Cho and colleagues (2009). Regions showing the main effects of relational complexity (shown in red), interference (shown in yellow; small volume corrected, uncorrected cluster-forming threshold $T > 2.3$, corrected cluster extent significance threshold, $p < .05$), and regions where

main effects overlapped (blue) within an a priori defined anatomical ROI mask of the bilateral MFG and IFG pars opercularis and pars triangularis. R, right; L, left. Coordinates are in MNI space (mm). (See color plate 83.)

to resolve interference during analogical reasoning. This finding is consistent with a broad role for these regions in cognitive control. By manipulating multiple cognitive demands simultaneously in a single reasoning task, the Cho and colleagues (2009) study was able to elucidate the neural architecture underlying behavioral interactions between complex cognitive processes. These findings show that analogical reasoning, which requires integration of multiple relations in the face of interference, is associated with the coordination of activity in multiple functionally dissociable regions of the prefrontal cortex. These subregions include those that are relatively more sensitive to demands on one component process, as well as regions that are jointly taxed by both relational integration and interference resolution.

THE TOPOGRAPHY OF THINKING As our survey of recent research makes clear, recent neuropsychological studies have demonstrated the dependence of relational reasoning on PFC, and neuroimaging studies have begun to delineate the functional anatomy of reasoning processes. These studies have converged on the finding that the frontopolar region is specifically activated when a problem requires the integration of multiple relations. As reviewed by Ramnani and Owen (2004), this region has several properties consistent with its playing an important role in complex cognitive tasks. First, the frontopolar region is comparatively larger in humans than in nonhuman primates. Second, unlike other PFC regions, it primarily has reciprocal connections with other supramodal regions in the PFC, suggesting that it is suited for processing abstract information. Third, the cellular

properties of the frontopolar cortex are consistent with its being a site of convergence. Although the density of neurons is not particularly high in this region, postsynaptic spines are extremely plentiful and dense in the dendritic arbors of these neurons.

While neuroimaging data from studies using reasoning tasks is consistent with the hypothesis that the frontopolar cortex subserves relational integration, this region is also activated in several other tasks involving complex cognition. For example, frontopolar cortex is active when subjects interrupt one task temporarily to perform another (a task requirement termed "branching"; Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999) and during episodic memory retrieval (Rugg et al., 1998). It may be possible to develop a theory of frontopolar cortex function that can account for the full range of findings. According to one proposal, the role of the frontopolar cortex is to process internally generated information (Christoff & Gabrieli, 2000). Relational integration might be subserved under this view if it is assumed that the reasoner must manipulate internally generated relations. On the face of it, however, the basic relations that must be integrated in a task such as the RPM (e.g., change in size or shape of a figure) are perceptually available in the problem itself. According to another view, the role of the frontopolar cortex is to integrate the results of multiple cognitive operations (Ramnani & Owen, 2004). Rather than focusing on the type of information being manipulated, this view emphasizes the cognitive processes supported by this region. Ramnani and Owen's proposal appears to be consistent with the relational integration hypothesis, as the outputs of cognitive processes underlying planning and problem solving (e.g., selected operators and new subgoals) may constitute relations that need to be integrated in order to sequence actions.

Another approach to developing a more precise theory of frontopolar functions is to search for subdivisions that support distinct cognitive processes. Recently, Wendelken and colleagues (2008) argued that the frontopolar region activated by relational integration in analogical reasoning is more lateral than the frontopolar region activated by branching, which is in the most rostral aspect of this region. A similar conclusion was reached by Gilbert and colleagues (2006) in a meta-analysis of neuroimaging studies showing frontopolar activations. Gilbert and colleagues also concluded that multitasking paradigms elicited more rostral activation. Thus it is possible that the most rostral subregion is involved in coordinating multiple task sets and goals, whereas more caudal, lateral subregions are involved in integration of relations (or perhaps, those relations represented in an explicit, declarative code of the sort employed in models of relational thought, such as LISA).

It also appears that there may be dorsoventral differences within the lateral frontopolar cortex, with more dorsal areas

activated during solution of visuospatial problems such as the RPM task and more ventral areas activated during verbal analogy tasks. Such variations may reflect a difference in the type of materials used in the tasks (visuospatial versus verbal), or other differences between the tasks (e.g., generating an analogical solution versus evaluating an analogical mapping). Future research using high-resolution fMRI may be able to further tease apart the contributions of distinct subregions of frontopolar cortex.

Over the past decade, understanding the neural basis of relational reasoning has for the first time become a tractable research problem. Much of the recent progress has been the product of neuroimaging and neuropsychological studies guided by theoretical frameworks developed in cognitive science. As the component processes of reasoning are mapped onto brain regions, it will be possible to delineate a functional network for reasoning. This accomplishment will represent a major step toward understanding the remarkable capabilities of human thought.

ACKNOWLEDGMENTS This chapter was written with support from the Office of Naval Research (Grant N000140810186).

REFERENCES

- ARON, A. R., ROBBINS, T. W., & POLDRACK, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.*, 8, 170–177.
- ASAAD, W. F., RAINER, G., & MILLER, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron*, 21, 1399–1407.
- BADRE, D., POLDRACK, R. A., PARE-BLAGOEV, E. J., INSLER, R. Z., & WAGNER, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, 47, 907–918.
- BRUN, A. (1993). Frontal lobe degeneration of non-Alzheimer's type revisited. *Dementia*, 4, 126–131.
- BUNGE, S. A., WENDELKEN, C., BADRE, D., & WAGNER, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cereb. Cortex*, 15, 239–249.
- CHO, S., HOLYOAK, K. J., & CANNON, T. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Mem. Cogn.*, 35, 1445–1455.
- CHO, S., MOODY, T. D., FERNANDINO, L., MUMFORD, J. A., POLDRACK, R. A., CANNON, T. D., KNOWLTON, B. J., & HOLYOAK, K. J. (2009). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. Manuscript under review.
- CHRISTOFF, K., & GABRIELI, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostro-caudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, 28, 168–186.
- CHRISTOFF, K., PRABHAKARAN, V., DORFMAN, J., ZHAO, Z., KROGER, J. K., HOLYOAK, K. J., & GABRIELI, J. D. E. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage*, 14, 1136–1149.
- D'ESPOSITO, M., POSTLE, B. R., JONIDES, J., & SMITH, E. E. (1999). The neural substrate and temporal dynamics of interference

- effects in working memory as revealed by event-related functional MRI. *Proc. Natl. Acad. Sci. USA*, 96, 7514–7519.
- DOUMAS, L. A. A., & HUMMEL, J. E. (2005). Approaches to modeling human mental representations: What works, what doesn't, and why. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 73–91). Cambridge, UK: Cambridge University Press.
- FALKENHAINER, B., FORBUS, K. D., & GENTNER, D. (1989). The structure-mapping engine: Algorithm and examples. *Artif. Intell.*, 41, 1–63.
- GENTNER, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cogn. Sci.*, 7, 155–170.
- GICK, M. L., & HOLYOAK, K. J. (1980). Analogical problem solving. *Cogn. Psychol.*, 12, 306–355.
- GIEDD, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Ann. NY Acad. Sci.*, 1021, 105–109.
- GILBERT, S. J., SPENGLER, S., SIMONS, J. S., STEELE, J. D., LAWRIE, S. M., FRITH, C. D., & BURGESS, P. W. (2006). Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. *J. Cogn. Neurosci.*, 18, 932–948.
- GREEN, A. E., FUGELSANG, J. A., KRAEMER, D. J., SHAMOSH, N. A., & DUNBAR, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Res.*, 1096, 125–137.
- HALFORD, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.
- HALFORD, G. S., WILSON, W. H., & PHILLIPS, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behav. Brain Sci.*, 21, 803–831; discussion, 831–864.
- HODGES, J. R., PATTERSON, K., OXBURY, S., & FUNNELL, E. (1996). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, 115, 1783–1806.
- HOLYOAK, K. J. (2005). Analogy. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 117–142). Cambridge, UK: Cambridge University Press.
- HOLYOAK, K. J., & THAGARD, P. (1989). Analogical mapping by constraint satisfaction. *Cogn. Sci.*, 13, 295–355.
- HUMMEL, J. E., & HOLYOAK, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychol. Rev.*, 104, 427–466.
- HUMMEL, J. E., & HOLYOAK, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychol. Rev.*, 110, 220–264.
- JONIDES, J., SMITH, E. E., MARSHUETZ, C., KOEPPE, R. A., & REUTER-LORENZ, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proc. Natl. Acad. Sci. USA*, 95, 8410–8413.
- KOECHLIN, E., BASSO, G., PIETRINI, P., PANZER, S., & GRAFMAN, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature*, 399, 148–151.
- KRAWCZYK, D. C., MORRISON, R. G., VISCONTAS, I., HOLYOAK, K. J., CHOW, T. W., MENDEZ, M. F., MILLER, B. L., & KNOWLTON, B. J. (2008). Distraction during relational reasoning: The role of prefrontal cortex in interference control. *Neuropsychologia*, 46, 2020–2032.
- KROGER, J. K., SAAB, F., 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. *Cereb. Cortex*, 12, 477–485.
- MILLER, E. K., & COHEN, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.*, 24, 167–202.
- MORRISON, R. G., DOUMAS, L. A. A., & RICHLAND, L. E. (2006). The development of analogical reasoning in children: A computational account. In R. Sun & N. Miyake (Eds.), *Proceedings of the Twenty-eighth Annual Conference of the Cognitive Science Society* (pp. 603–608). Mahwah, NJ: Erlbaum.
- MORRISON, R. G., HOLYOAK, K. J., & TRUONG, B. (2001). Working memory modularity in analogical reasoning. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society* (pp. 663–668). Mahwah, NJ: Erlbaum.
- MORRISON, R. G., KRAWCZYK, D. C., HOLYOAK, K. J., HUMMEL, J. E., CHOW, T. W., MILLER, B. L., & KNOWLTON, B. J. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *J. Cogn. Neurosci.*, 16, 260–271.
- NEE, D. E., WAGER, T. D., & JONIDES, J. (2007). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cogn. Affect. Behav. Neurosci.*, 7, 1–17.
- PENN, D. C., HOLYOAK, K. J., & POVINELLI, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and non-human minds. *Behav. Brain Sci.*, 31, 109–178.
- PRAHAKARAN, V., SMITH, J. A. L., DESMOND, J. E., GLOVER, G. H., & GABRIELI, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cogn. Psychol.*, 33, 43–63.
- RAMNANI, N., & OWEN, A. M. (2004). Anterior prefrontal cortex: Insights from anatomy and neuroimaging. *Nat. Rev. Neurosci.*, 5, 184–194.
- RAVEN, J. C. (1941). Standardization of progressive matrices, 1938. *Br. J. Med. Psychol.*, 19, 137–150.
- RAZ, N., GUNNING, F. M., HEAD, D., DUPIUIS, J. H., MCQUAIN, J., BRIGGS, S. D., LOKEN, W. J., THORNTON, A. E., & ACKER, J. D. (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cereb. Cortex*, 7, 268–282.
- RICHLAND, L. E., MORRISON, R. G., & HOLYOAK, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *J. Exp. Child Psychol.*, 94, 249–271.
- 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.
- RUGG, M. D., FLETCHER, P. C., ALLAN, K., FRITH, C. D., FRAKOWIAK, R. S., & DOLAN, R. J. (1998). Neural correlates of memory retrieval during recognition memory and cued recall. *NeuroImage*, 8, 262–273.
- SHIMAMURA, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, 28, 207–218.
- SNOWDEN, J., NEARY, D., & MANN, D. (2007). Frontotemporal lobar degeneration: Clinical and pathological relationships. *Acta Neuropathol.*, 114, 31–38.
- STERNBERG, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.
- STERNBERG, R. J., & NIGRO, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Dev.*, 5, 127–138.
- THAGARD, P. (1996). The concept of disease: Structure and change. *Commun. Cogn.*, 29, 445–478.
- THOMPSON-SCHILL, S. L., D'ESPOSITO, M., AGUIRRE, G. K., & FARAH, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc. Natl. Acad. Sci. USA*, 94, 14792–14797.
- THOMPSON-SCHILL, S. L., JONIDES, J., MARSHUETZ, C., SMITH, E. E., D'ESPOSITO, M., KAN, I. P., KNIGHT, R. T., & SWICK, D. (2002).

- Effects of frontal lobe damage on interference effects in working memory. *Cogn. Affect. Behav. Neurosci.*, 2, 109–120.
- TOHILL, J. M., & HOLYOAK, K. J. (2000). The impact of anxiety on analogical reasoning. *Thinking & Reasoning*, 6, 27–40.
- VISKONTAS, I. V., HOLYOAK, K. J., & KNOWLTON, B. J. (2005). Relational integration in older adults. *Thinking & Reasoning*, 11, 390–410.
- VISKONTAS, I. V., MORRISON, R. G., HOLYOAK, K. J., HUMMEL, J. E., & KNOWLTON, B. J. (2004). Relational integration, inhibition and analogical reasoning in older adults. *Psychol. Aging*, 19, 581–591.
- WAGNER, A. D., PARE-BLAGOEV, E. J., CLARK, J., & POLDRACK, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31, 329–338.
- WALTZ, J. A., KNOWLTON, B. J., HOLYOAK, K. J., BOONE, K. B., BACK-MADRUGA, C., MCPHERSON, S., MASTERMAN, D., CHOW, T., CUMMINGS, J. L., & MILLER, B. L. (2004). Relational integration and executive function in Alzheimer's disease. *Neuropsychology*, 18, 296–305.
- WALTZ, J. A., KNOWLTON, B. J., HOLYOAK, K. J., BOONE, K. B., MISHKIN, F. S., SANTOS, M. D., THOMAS, C. R., & MILLER, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychol. Sci.*, 10, 119–125.
- WALTZ, J. A., LAU, A., GREWAL, S. K., & HOLYOAK, K. J. (2000). The role of working memory in analogical mapping. *Mem. Cogn.*, 28, 1205–1212.
- WENDELKEN, C., NAKAHBENKO, 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. *J. Cogn. Neurosci.*, 20, 682–693.
- WHARTON, C. M., GRAFMAN, J., FLITMAN, S. S., HANSEN, E. K., BRAUNER, J., MARKS, A., & HONDA, M. (2000). Toward neuro-anatomical models of analogy: A positron emission tomography study of analogical mapping. *Cogn. Psychol.*, 40, 173–197.
- ZHANG, J. X., FENG, C. M., FOX, P. T., GAO, J. H., & TAN, L. H. (2004). Is left inferior frontal gyrus a general mechanism for selection? *NeuroImage*, 23, 596–603.