Full Frontal Fluidity?

Looking in on the Neuroimaging of Reasoning and Intelligence

Michael J. Kane

The centres of inhibition being thus the essential factor of attention, constitute the organic basis of all the higher intellectual faculties.

Ferrier (1876, p. 287)

... the frontal lobes... with their associated sensory centres, form the substrata of those psychical processes which lie at the foundation of the higher intellectual operations.

Ferrier (1886, p. 467)

Is there a biological seat of reason—a "g-gram," or Cartesian pineal gland—for intelligence? David Ferrier (1876, 1886) thought so, and he saw it in the brain's frontal lobe. His idea was that complex thought, in the form of sensory-motor associations, occurred while motor output was inhibited via attential processes originating in the prefrontal cortex (PFC). Although the early 20th-century zeitgeist was divided over the PFC's contribution to higher mental functions, based largely on the case studies of patients with brain damage (e.g., Goldstein, 1944; Hebb, 1945), Ferrier's localization of attention and intelligence to the PFC seems prescient today.

In a strict sense, of course, Ferrier (1876, 1886) must have been wrong. As any of the contributors to this volume can attest, intelligence is a multifaceted construct if it is a single construct at all, and neuroscientists have long known that the brain is a massively interconnected system, in which many specialized regions contribute to performance of even the simplest cognitive tasks. Thus, no single brain area is completely responsible for the performance of any one cognitive task, never mind accounting for a broad cognitive construct. Indeed, no specific locus of brain damage or brain activity has proven to be uniquely or consistently associated with performance on standardized intelligence test batteries.

Despite these facts, however, there is reason to believe that at least one aspect of intelligence—general fluid ability (Gf)—is
closely linked to structures of the PFC. Gf is a statistical construct, derived from the factor analysis of novel reasoning tests, and it suggests a general ability to reason through problems that afford little benefit of domain-specific knowledge or skill. That is, Gf reflects the fact that diverse tests of novel reasoning correlate substantially with one another, indicating that people who perform well on one reasoning test also tend to perform well on another. Of importance here, these tests of novel fluid reasoning, which make substantial working memory and attention control demands (e.g., Carpenter, Just, & Shell, 1990; Embrerton, 1995; see also Chapter 21, this volume), are selectively impaired by PFC damage. These Gf tests also appear to selectively activate particular PFC regions and circuitry in healthy humans, as indicated by in vivo neuroimaging studies.

In this chapter, I first provide a bird’s-eye view of the cognitive psychological research suggesting behavioral links among working memory capacity, executive attention, and intellectual ability, followed by an overview of the neuroscience evidence linking working memory capacity and executive attention functions to PFC. I then critically and comprehensively review the evidence from modern neuroimaging research suggesting that the PFC plays an important role in Gf. That is, I consider findings from studies that have measured participants’ regional cerebral blood flow via positron emission tomography (PET) or single photon emission computerized tomography (SPECT) or regional blood oxygenation via functional magnetic resonance imaging (fMRI) during the performance of fluid reasoning tasks. To preview, I argue that the preponderance of neuroimaging evidence shows that several PFC regions, particularly on the lateral surface, are critically involved in the performance of particular Gf tests. However, if one takes the view that Gf is a multivariate, individual-differences construct, representing the performance variance that is shared among a variety of novel reasoning tasks, then neuroimaging research has so far told us quite little about the relevance of the PFC to Gf and individual differences.

**Working Memory Capacity, Executive Attention, and Gf**

Baddeley and Hitch (1974) first proposed the working memory (WM) construct as a means to capture the functional importance of immediate memory. Their idea was that rote rehearsal is an unlikely function for an evolutionarily adaptive memory system and that short-term retention should be most beneficial in maintaining privileged access to task-relevant information in the context of ongoing mental work. That is, immediate memory should serve complex cognition by keeping critical stimuli, ideas, or response productions accessible enough to influence performance. Immediate memory is therefore a “working” memory. Baddeley and Hitch (1974; Baddeley, 1986) proposed a WM model in which domain-specific memory buffers interacted with a general attentional resource, or control structure (the “central executive”), and the memory and attention components of this WM system worked together to integrate ongoing processing and storage operations.

Daneman and Carpenter (1980) subsequently created an individual-differences measure of WM capacity, “reading span,” in which participants remembered unrelated words while simultaneously processing sentences for meaning. Participants read aloud or listened to series of two to six sentences, after which they recalled the final word from each sentence. Thus, a traditional span task of short-term memory (STM) was embedded within a secondary, background task of reading, thereby engaging executive attention as well as storage processes. Subsequent research has developed a variety of these WM tasks, pairing processing demands such as verifying equations, counting shapes, or mentally rotating letters, with storage demands of retaining a series of words, letters, numbers, or visuospatial figures (e.g., Case, Kurland, & Goldberg, 1982; Shah & Miyake, 1996; Turner & Engle, 1989). What is most interesting about these WM span tasks is that, in contrast to traditional STM tasks involving rote repetition of digits or words, they consistently predict a broad range of cognitive abilities. A large
literature now clearly demonstrates that WM span scores correlate substantially with measures of comprehension, learning, communicating, and reasoning, virtually independently of the particular stimuli used in the WM tasks (for reviews, see Daneman & Merikle, 1996; Engle, 1998, 2001).

Why do WM span tasks work so well as individual-differences measures? My colleagues and I have argued that it is not due to their memory storage demands per se (Engle, 2002; Engle & Kane, 2004; Engle, Kane, & Tuholski, 1999; Kane & Engle, 2002; see also Chapters 4, 5, and 21, this volume). If storage were critical, then simple STM span tasks should predict cognitive abilities better than they do. Instead, what drives the greater predictive power of WM span tasks is that they tap the attention control ability to maintain access to information in the face of interference and distraction. WM span tasks require participants to keep target stimuli in mind under proactive interference from similar items on prior trials (e.g., May, Hasher, & Kane, 1999), and keeping those stimuli in mind is challenged by intermittently shifting attentional focus between the memory and secondary-processing tasks (e.g., Hitch, Towsse, & Hutton, 2001). Because proactive interference impairs retrieval from long-term memory, it forces participants to rely on the sustained, active maintenance of the target information instead. Such maintenance is threatened, however, by the secondary-processing task preventing participants from keeping the target items in attentional focus. We claim that executive control processes help keep target items maintained and accessible in the face of such interference and attention shifts. Moreover, this domain-general executive attention capability is responsible for the broad predictive utility of WM span tasks.

Although the predictive utility of WM span tasks alone suggests that they tap into a domain-general capability such as Gf, recent research has established this link more directly. The critical hypothesis is, of course, that WM capacity, with its attendant executive processes, is a limiting factor in complex cognitive processing. Therefore, individual differences in WM span tasks should account for a significant proportion of the variation in Gf. In retrospect, this hypothesis gained initial support from Wittenborn’s (1943) collection of psychometric “attention” tasks. These tasks required simultaneous processing and storage of information and comprised a statistically reliable factor that was closely related to psychometric Gf (see also Stankov, 1983). More recently, Kyllonen and Christal (1990) tested more than 2,000 participants in various WM tasks, some similar to Wittenborn’s, across four confirmatory factor-analytic studies. Latent variables derived from the shared variance among these tasks correlated between .80 and .90, with latent variables comprising verbal and nonverbal reasoning tasks (see also Kyllonen, 1993; but for estimates of WM/Gf correlations in the .50–.70 range, see Ackerman, Beier, & Boyle, 2002; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002).

To test the idea that the attention and not storage demands of WM tasks drive their broad predictive power, latent variable studies by Engle, Tuholski, Laughlin, and Conway (1999) and Conway, Cowan, Bunting, Therriault, and Minkoff (2002) contrasted relations among WM, STM, and Gf tasks, with the latter indexed by Raven’s Standard Progressive Matrices (Raven, Raven, & Court, 1998) and Cattell’s Culture Fair tests (Institute for Personality and Ability Testing, 1973). Confirmatory factor analyses indicated that factors derived from the WM and STM tasks were separable but correlated, suggesting that WM span tasks capture something, perhaps attentional, beyond just simple storage. Structural equation models tested the critical attention hypothesis: After removing the variance common to the latent variables for WM and STM, which should represent their shared storage requirements, the residual variance in the WM factor should correlate with Gf. This residual variance from WM, after partialling out the storage variance associated with STM storage, should reflect the unique executive attention demands of WM tasks. In fact, both studies found that the WM residual accounted for approximately 25% to 35% of the variance in Gf, whereas the shared variance with STM accounted for 9% or less (the residual variance from STM did not correlate with Gf). The critical factor in the relation...
between WM capacity and fluid intelligence thus appears to be executive attention. More direct evidence linking WM capacity to executive attention has come from quasi-experimental studies in which high and low WM span participants (i.e., university students at the top and bottom quartiles of WM span scorers) were tested in attention tasks. For example, Kane and Engle (2000) tested participants in a Brown-Peterson task with three 10-word lists drawn from the same semantic category. Proactive interference is typically seen in such preparations, with recall dropping for each subsequently presented list as prior items block or intrude into recall. In this study, some participants had their attention divided during study or recall by having to continuously tap a pseudo-random finger sequence. The logic was that if resisting proactive interference requires attention control, then interference vulnerability should vary with WM capacity and divided-attention demands. Indeed, in single-task conditions, WM span groups did not differ in recall of List 1, but low-span individuals suffered significantly greater proactive interference on List 3 than did high spans, with recall dropping by 50% and 30%, respectively. Moreover, under dual-task conditions, high spans became as susceptible to interference as low spans, but low spans’ susceptibility did not change. High spans, under divided attention, thus simulated normal low-span behavior! High WM appears to be associated with an executive attention capability to block interference that may be disrupted by dual tasking. Moreover, the fact that low spans’ interference effects did not change under divided attention suggest that they normally did not engage executive processes to combat interference: without using controlled processing in the first place, their poorer performance could not be reduced by dividing attention.

Consistent with an attentional view of WM capacity, high- and low-span individuals also differ in relatively “simple” attention tasks that make few demands on memory retrieval. For example, in a dichotic listening task in which participants’ names were spoken once in the nonshadowed distractor message, low-span participants were more than three times more likely to report hearing their names than were high spans (65% vs. 20%, respectively: Conway, Cowan, & Bunting, 2001). Parallel findings come from visual attention tasks, such as the anti-saccade (Kane, Bleckley, Conway, & Engle, 2001) and Stroop tasks (Kane & Engle, 2003). In the anti-saccade task, participants must look either toward or away from a flashing cue in the periphery. Span groups did not differ when required to look toward the flashing cue in “pro-saccade” task blocks, in which performance was guided by automatic orienting. However, when required to look away from the cue in “anti-saccade” blocks, low spans made more eye movement errors, corrected these errors more slowly, and made correct eye movements more slowly than did high spans. In other words, low WM individuals were impaired in blocking automatic responses in favor of novel goal-directed ones. Finally, low spans are significantly more vulnerable to Stroop interference than are high spans. When participants must ignore the meaning of color words and instead report the colors in which they appear, low spans either make more errors or respond more slowly than high spans, depending on the task context. Low spans are particularly error prone when a majority of Stroop trials are congruent, presenting words in their matching colors, and when controlled processing is therefore unnecessary for correct responding. In these contexts, low spans appear to lose access to the novel goal of ignoring the words and instead “zone out” and begin reading the words. We view all of these related WM span differences as reflecting variability in executive attention—that is, in the ability to maintain access to novel task goals in the face of memory interference and salient conflict in the environment.

It is probably worth noting that more direct relations between attention control and Gf have been explored and found outside of the WM tradition. For example, Duncan, Williams, Nimmo-Smith, and Brown (1993) found that some component skills of automobile driving correlated more strongly with scores on the Culture Fair test of Gf than did others (e.g., visual scanning). Most interestingly, the component skills that shared the most variance with Gf were also the ones most susceptible to interference from the secondary task of verbally producing a continuous, random string of digits. The correlations between the Gf loadings of
component skills and performance loss under dual-task conditions were robust and ranged from $r = .57$ to $.67$.

Similarly, when participants must combine auditory memory tasks with other, unrelated cognitive tasks, the correlation between the tasks is higher than the correlation between these same tasks when performed singly, suggesting individual differences in the ability to maintain performance among multiple simultaneous tasks (Fogarty & Stankov, 1982; Stankov, 1983). Performance on these combined tasks also correlates highly with Gf measures and significantly more highly than do scores on these tasks performed alone (Roberts, Beh, & Stankov, 1988). In a study that tied these dual-task abilities back to notions of WM capacity, Splishby (1992) had participants perform two main tasks, either singly or simultaneously. In one of the tasks, participants heard a sequence of sounds and kept a mental tally of four different sound types. In the second task, participants took a visually presented vocabulary test. Performance on the dual task of auditory tallying plus vocabulary correlated significantly with a Gf composite derived from an additional test battery. Furthermore, the dual-task correlation with Gf was significantly stronger than the Gf correlations with either of the tasks performed individually, as previously reported. Thus, time-sharing between tasks was more closely correlated with Gf than was the performance of either task alone. In addition to these main tasks used to assess divided attention, Splishby included two WM tests from the Wittenborn (1943) battery, and she found that partiailling out WM reduced the dual-task/Gf correlation substantially. Thus, WM capacity shares significant variance with attention control and Gf capabilities.

**Working Memory, Executive Attention, and the Lateral Prefrontal Cortex**

On one hand, the search for the intellectual functions of the PFC has followed clinical observations of patients with PFC injury, who display a variety of both subtle and serious difficulties in complex cognitive tasks in the laboratory and in life (e.g., Lezak, 1983; Luria, 1966; Shallice & Burgess, 1991). A complication that arises from this line of work, however, is that patients with PFC damage often show normal psychometric IQ (e.g., Eslinger & Damasio, 1985; Hebb & Penfield, 1940; Weinstein & Teuber, 1957), a point to which I will later return. On the other hand, the link between PFC and Gf has also been inferred from the observed reliance on PFC circuitry of WM and executive attention functions (see Kane & Engle, 2002). Given the established behavioral associations among WM, attention, and Gf, it seems reasonable that these constructs might share a neural substrate.

It is worth noting that, by arguing for a special role for lateral prefrontal cortex (LPFC) areas in WM and attention control capacities, I do not claim that the PFC solves these cognitive problems in isolation. The LPFC, by itself, does not embody WM and executive attention functions. Instead, PFC regions, including lateral ones, are richly interconnected with other cortical and subcortical structures, as well as with each other (see, e.g., Barbas & Mesulam, 1985), and so these networked regions undoubtedly contribute to the tasks and functions to be reviewed. That said, we shall also see that LPFC cells do appear to have important properties that distinguish them from other neurons.

The WM function of PFC cells was first demonstrated through the delayed-memory experiments of Jacobsen (1935, 1936), in which macaque monkeys viewed two wells, one of which was baited with food on each trial, and the monkey had to remember the baited well over some filled delay period. These studies have been extensively replicated and expanded on by modern researchers (see Fuster, 1980; Goldman-Rakic, 1987), who have found, in short, that permanent or temporary damage to cells in the LPFC bring delayed-memory performance to chance levels, even with delays of only a few seconds.

In addition, single-cell recordings find that individual cells in the LPFC begin firing during the encoding phase of trials that feature their preferred type of target (e.g., features in a particular region of space), and they maintain their activity throughout the entire delay period. Some LPFC cells appear to fire selectively for particular stimulus locations, others for particular stimulus objects, and still others for particular rules or even combinations of any of
the above (e.g., Rainer, Asaad, & Miller, 1998). Indeed, as a whole, LPFC cells appear to be domain general in their active maintenance of memory representations, interacting with different posterior networks depending on the stimuli and task at hand. Moreover, unlike delay-sensitive cells in more posterior cortical areas, LPFC cells maintain their delay-period activity even when distracting objects intervene between target encoding and test (e.g., Miller & Desimone, 1994; Miller, Erickson, & Desimone, 1996). Thus, Kane and Engle (2002) suggested that the PFC is likely to be a critical structure in the executive maintenance of information in the presence of interference and distraction, the putative basis of a domain-general WM capacity (see also Malmo, 1942; Miller & Cohen, 2001).

Most human neuropsychological studies show similar results, with patients with PFC damage demonstrating greater difficulties with delayed-memory tasks than do patients with more posterior brain damage (e.g., Bechara, Damasio, Tranel, & Anderson, 1998; Chorover & Cole, 1966). Neuroimaging studies with healthy adults also support the link between WM and PFC. LPFC regions are activated over baseline in both delayed-memory tasks (e.g., D’Esposito, Ballard, Zarahn, & Aguirre, 2000; McIntosh, Grady, Haxby, Ungerleider, & Horwitz, 1996) and n-back tasks, in which participants report whether each stimulus in a long sequence matches the stimulus n items back (typically two or three back; e.g., Braver et al., 1997; Cohen et al., 1994). Moreover, unlike the WM-related activity in many posterior brain areas, the PFC activity increases lawfully with memory load and stimulus manipulation demands.

LPFC structures are clearly critical to WM maintenance. Of importance, evidence converges from attention research to indicate a role for LPFC regions in executive control, as well. For example, several human imaging studies show that LPFC is particularly activated during dual tasks. In perhaps the best-known such study, D’Esposito et al. (1995) tested participants in two different tasks, one verbal and one spatial, which were performed alone and together. By themselves, neither task elicited LPFC activation, even when the spatial task was speeded to match its difficulty to the dual task. However, when the two tasks were performed together, LPFC was highly activated. LPFC areas are also consistently activated in tasks demanding that a novel goal be maintained in the face of conflict, such as in anti-saccade tasks (e.g., Sweeney et al., 1996) and Stroop tasks (e.g., Zysset, Müller, Lohman, & von Cramon, 2001). In a Stroop task that cued whether each trial was to be an incongruent or a congruent trial, LPFC activation rose steadily over the 11 seconds following the cue on interference trials only, suggesting that the activation was related to establishing and maintaining the ignore-the-word set (MacDonald, Cohen, Stenger, & Carter, 2000). Indeed, the strength of this goal activation was negatively correlated with the interference score ($r = -0.63$), such that participants showing greater PFC activation showed less difficulty in naming the color and ignoring the word. Together, these findings suggest that LPFC activity is important to maintaining task goals in preparation for conflict, in line with our findings of WM capacity differences in anti-saccade and Stroop tasks. Indeed, human patients with LPFC damage perform quite similarly to low WM participants in these tasks, albeit with a greater impairment (e.g., Perret, 1974; Pierrot-Deseilligny, Rivaud, Gaymard, & Agid, 1991).

To summarize, behavioral research from an individual-differences perspective suggests that WM capacity, executive attention, and Gf are strongly interrelated. Although neither WM capacity nor executive attention measures correlate strongly enough with Gf to indicate an isomorphism (see Conway, Kane, & Engle, 2003), WM and attention control may be among a few cognitive abilities that underlie general fluid abilities. The fact that LPFC areas also seem particularly important to active maintenance and control processes, across a wide range of tasks and stimulus domains, also lends support to the idea that these constructs are linked. I turn now to reviewing the neuroscience evidence that LPFC structures play a central role in individual differences and Gf.

**A Front Seat for Reason: Neuropsychological Evidence**

As mentioned earlier, a difficulty for the idea that LPFC regions are important to fluid
intelligence is the fact that patients with frontal lobe damage often have normal intelligence, as measured by psychometric IQ batteries. Of course, there are a number of obstacles to interpreting such neuropsychological findings. First, measures of premorbid IQ are rarely available. If LPFC damage limits attention control, then it should affect individuals of high intelligence more substantially than those of lower intelligence (Kane & Engle, 2002). This would be analogous to Kane and Engle’s (2000) empirical finding that dividing high spans’ attention simulated low-span vulnerability to memory interference, but dividing low spans’ attention had little effect. Second, frontal damage often resulted from surgery designed to correct epilepsy or psychosis, premorbid diseases that almost certainly had lasting effects on general cognitive functioning. Moreover, when LPFC damage resulted from strokes or tumor surgery, it often did so in relatively old adults, who, along with their controls, already suffer losses in Gf and so may not show obvious signs of postinjury deterioration. Third, these points about etiology make clear that, in the literature, patients with prefrontal damage suffered damage via many different sources, including surgery, strokes, penetrating head wounds, and closed head injuries, and so comparing across cases is quite difficult, if not misleading. Fourth, brain damage is often diffuse and encroaches on multiple areas of brain tissue (not to mention vascular systems), and so prefrontal damage is often attended by more posterior damage and vice versa.

Duncan (1995) has raised yet a fifth complication, but one that may resolve some of the confusion in the literature and also lead to testable predictions. Namely, psychometric IQ batteries yield an estimate of general intelligence by averaging across numerous subtests, only some of which require novel fluid reasoning and many of which require acquired knowledge, cognitive skills, and strategies. If LPFC structures are particularly important for fluid aspects of reasoning, as opposed to more crystallized abilities, then psychometric intelligence batteries may not be sensitive to PFC-related impairment. Indeed, Duncan has amassed some new evidence to support his claim. Duncan, Burgess, and Emslie (1995) tested three patients with localized frontal damage and healthy controls matched on age and psychometric IQ (via the Wechsler Adult Intelligence Scale [WAIS] or the Wechsler Adult Intelligence Scale–Revised [WAIS-R]). All IQ scores were quite high, ranging from 126 to 130. These participants were then compared on a nonverbal Gf test, the Cattell Culture Fair; as predicted, the patients with frontal lobe damage performed much more poorly on the Culture Fair test than on the IQ batteries and much more poorly than their matched controls. For the patients with frontal lobe damage, their estimated IQ dropped more than 20 points from the WAIS to the Gf tests, and their Gf scores were more than 20 points below those of the controls. A second group of patients, with posterior brain damage, showed no WAIS-Gf discrepancy and no deficits relative to matched controls. Duncan, Emslie, Williams, Johnson, and Freer (1996) subsequently replicated this entire pattern of findings.

A Front Seat for Reason: Neuroimaging Evidence

Duncan’s (1995) hypothesis about PFC and Gf has held up to empirical test, but the database is small, and few patients have been tested in each study. Moreover, as far as I know, there are actually more examples in the literature of patients with PFC damage performing normally on Gf tests than there are showing impairment (e.g., Berker & Smith, 1988; Brazzelli, Colombo, Della Sala, & Spinnler, 1994; McFie & Piercy, 1953; but see Waltz et al., 1999, for PFC-related impairments). Duncan’s ideas do not yet seem to have provided this literature’s panacea, and so further neuropsychological research is necessary to clarify the role of PFC in fluid ability.

An Introduction to the Imaging of Gf

Fortunately, converging evidence for or against the LPFC’s relation to Gf can be provided by in vivo neuroimaging techniques that measure, with varying degrees of spatial and temporal resolution, the activity of different brain regions during healthy participants’ performance of cognitive tasks. The first such study
of brain activity during Gf-loaded tasks appears to have been conducted by Risberg and Ingvar (1973), who tested performance on Raven’s Standard Progressive Matrices (RSPM) and a figural “odd-man-out” task (Dureman & Sälde, 1959) in 12 hospital inpatients (with 2 completing the RSPM and 10 completing the odd-man-out task) using the 133Xenon injection technique. The state of the art at the time of Risberg and Ingvar’s publication, the 133Xenon injection technique required a small amount of a radio-active isotope to be injected into the participant’s carotid artery, with the needle left in place throughout the entire test session to track regional cerebral blood flow (rCBF). Although 133Xenon injection was considerably more invasive than are more modern neuroimaging methods, more important from a scientific perspective is the fact that the spatial and temporal resolution of the resulting brain scans was very poor. In Risberg and Ingvar’s study, the spatial resolution was approximately 20 mm, in contrast to the 2- to 4-mm resolution of current fMRI methods. The temporal resolution was 15 minutes, with many of the blood flow indices that the authors took dominated by hemodynamic events occurring over 2 to 6 minutes (in contrast to the 1- to 3-second resolution of current event-related fMRI methods).

What did Risberg and Ingvar (1973) find, given these limitations? Compared to a brain-wide rCBF average taken during “rest,” in which participants closed their eyes during scanning, the RSPM and odd-man-out tests activated frontal, pre-Rolandic, and parietal cortices, with the odd-man-out test recruiting occipitotemporal areas as well. Although it is difficult to ascertain quantitative differences in activation among these areas from the data figures, PFC areas appear to have been more strongly activated by the Gf tests than were most of the other cortical areas, with the exception of the occipitotemporal cortex. Much of this posterior activity probably reflects visual activity during the test (compared to rest). Also, given the relative activity of frontal regions during rest here (10% above the hemispheric mean), the test-minus-rest difference score favored the posterior brain regions. Even so, and despite the obvious limitations to Risberg and Ingvar’s methods, this seminal study points to the PFC as being at least one potentially important brain region to the performance of novel fluid-reasoning tasks.

An Introductory Critique of the Imaging of Gf

The past 30 years have seen an increasing interest in the neuroimaging of intelligence, and I will attempt to exhaustively review the findings from these studies below. By far, the most common empirical strategy in this literature has followed closely from Risberg and Ingvar (1973). Participants are scanned during performance of a putative Gf task, such as the RSPM, and brain activity during test is compared to (via subtraction) either a resting-state baseline or a control task that shares input and output features with the Gf task but not its reasoning demand. My review will begin here, considering separately lines of work investigating the performance of either Raven’s-like, nonverbal reasoning tasks (primarily inductive) or “three-term” verbal reasoning tasks in which participants must consider the validity or likelihood of a conclusion following from two premises (e.g., syllogistic-deductive and relational-inductive reasoning).

Although my conclusion will be that both varieties of the fluid-reasoning task clearly and reliably engage lateral prefrontal regions of the brain, I believe that these repeated demonstrations of LPFC activation during reasoning are of limited and decreasing scientific value, at least with respect to understanding the Gf construct. (In fairness, the studies of verbal reasoning were not explicitly directed at questions about Gf, despite their secondary relevance to this issue.) We can now say with some confidence that brain activity is variously localized during the performance of a given Gf test. However, we currently know virtually nothing about whether such neural activity has anything to do with individual differences in the performance of these tests or with the variance that is shared among Gf marker tests. If Gf is a construct reflecting the shared, individual-differences variance in performance among novel reasoning tasks, then Gf must be studied as a multivariate construct and not in isolation.

Consider that most automobiles rely on a similar braking system, the critical components
of which can be identified through "in vivo" observations of a stopping car. These components are analogous to those LPFCs and other cortical structures that are activated by performance of Gf marker tests. Yet, if we consider braking distance as a measure of brake system effectiveness, then the observed differences in braking distance among individual automobiles or makes of automobiles need not result only from variation in the performance of the brake pads, drums, and master cylinders. Differences might also—or instead—derive from variability in tire balance, the aerodynamics and weight of the chassis, or the ergonomic placement of the brake relative to the accelerator pedal, despite the fact that these are not basic mechanisms of braking. From this view, then, LPFC and other brain areas that are selectively engaged by individual tests of Gf (which, I should note, are not themselves perfect reflections of the Gf construct) may not be the structures responsible for individual differences in performance of those tests. These areas may not be the structures that are responsible for the covariation of individual differences among several such tests either. Indeed, given that most of the neuroimaging studies to date have tested fewer than 15 participants each, it is probably unwise to hold out too much hope of their elucidating the biology of cognitive individual differences.

Fortunately, the consistency of LPFC activation across individual studies (in contrast to the more erratic patterns from all other brain regions) provides indirect evidence that LPFC structures contribute to domain-general Gf. However, to more directly demonstrate the importance of the LPFC to individual differences, investigators must either correlate brain activity to performance within a single Gf task, directly compare brain activity and performance across multiple Gf tasks, or both. Given the centrality of individual differences to questions of Gf, surprisingly few investigations have taken these approaches (although such studies are quite expensive to conduct). My review concludes by considering what these few individual-differences studies have told us about the LPFC and Gf and what this research might suggest to be important avenues for future work.

As a preliminary caveat, let me note that the focus of this chapter is on the putative role of PFC regions in general fluid ability, and so in the interest of space and clarity, my review generally does not discuss the non-PFC areas that are selectively activated by Gf tests. Although this strategy might be argued to overestimate the importance of PFC regions relative to others, there are, in fact, no other brain regions that are consistently activated over baseline across the entire range of tasks considered here. So although particular posterior brain regions may be individually important for one putative Gf task or another (e.g., superior and inferior parietal cortex in the Raven's Progressive Matrices [RPM] test), none of these posterior structures is involved across the board. LPFC is.

Basic Prefrontal Mechanisms of Fluid Reasoning

The relevant neuroimaging literature may be organized according to the broad category of Gf tasks that have been investigated. I begin by reviewing studies using nonverbal, figural reasoning tasks such as the Raven's test, as these tasks may be the best individual markers of general fluid ability (Marshalek, Lohman, & Snow, 1983; see also Chapter 21, this volume). In general, such tasks present participants with an incomplete series or matrix of novel figures, and the participants' task is to induce the rule that governs the arrangement of figures and accordingly select the correct completion from among several foils. These tests often proceed from more easy to more difficult items, with the more difficult items characterized by the requirement to simultaneously consider and maintain more stimulus dimensions and apply more transformational rules (e.g., Carpenter et al., 1990; Embleton, 1995).

I will subsequently review a program of Gf studies using verbal materials, in which participants are presented with two premises—such as "All generals are men: few men enjoy knitting"—and must make a judgment about a conclusion that follows the premises, such as "No generals enjoy knitting." Although these tests are verbal, they obviously require fluid reasoning over and above contributions of crystallized knowledge. Some imaging studies have required participants to judge the logical validity of the conclusion via formal, deductive reasoning (the
example here is invalid). Others have asked participants to judge whether the conclusion is more likely to be literally true than false (the example here may be considered likely by many people). Additional interesting manipulations from this literature that I will consider involve the specific content of the premises and conclusions, as well as the kinds of relations that must be considered among the propositions. Some have been abstract (e.g., All Ps are Bs. No Bs are Ks. . . . No Ps are Ks.), some have presented content that conflicts with participants’ knowledge (e.g., All pets are poodles. All poodles are dangerous. . . . No pets are dangerous.), some have presented spatial relations among subjects (e.g., Officers stand next to generals. Privates stand behind generals. . . . Privates stand behind officers.), and some have presented nonspatial relations among subjects (e.g., Officers are heavier than generals. Generals are heavier than privates. . . . Privates are heavier than officers.). In most of these instances, LPFC is engaged over baseline.

Inductive Figural/Analytic Reasoning

Several rCBF-imaging studies followed Risberg and Ingvar (1973) in exploring the neural underpinnings of the RPM test, either the Standard (RSPM) version or the Advanced (RAPM). Results were surprisingly mixed. On the positive side, Risberg, Maximilian, and Prohovnik (1977) found the RAPM to activate frontal cortex very broadly (as a function of the spatial resolution) compared to a rest control in 12 healthy adults. Also, under PET. Esposito, Kirkby, Van Horn. Ellmore, and Berman (1999) contrasted the RSPM to a control task in which RPM-like items were presented but with all the figures identical and matching one of the response choices. Here, in 20 adults, the RSPM activated LPFC areas bilaterally (Brodmann Areas 9, 46). On the negative side, however, three older studies failed to find PFC areas activated over baseline during the RAPM; instead, only temporal, parietal, and occipital areas were engaged (Berman, Illowsky, & Weinberger. 1988; Haier et al., 1988; Maximilian, Prohovnik, & Risberg, 1980). I would explain these null findings as being at least partly the result of very poor spatial resolution, with Berman et al. (1988) averaging all cortical rCBF into only five large brain regions and Haier et al. (1988) dividing three transverse brain slices into only eight pie-shaped wedges each, leaving questionable sensitivity to pick up a uniquely lateral contribution of PFC to performance. In addition, the Maximilian et al. (1980) result is suspect due to the very high level of frontal activation during the baseline rest condition, thus making the deck against finding an above-baseline PFC effect.

Fortunately, the impact of these problematic negative results is diminished by three more recent studies, using fMRI with excellent spatial and temporal resolution, which investigated interesting variations on the basic RPM test. All of these studies contrasted figural reasoning to control tasks matched on perceptual and response requirements, and all manipulated the cognitive demand of the task across items. As illustrated in Figure 9.1, Prabhakaran, Smith, Desmond, Glover, and Gabrieli (1997) adapted the RPM to present some items requiring primarily figural/spatial reasoning (Panel B). with only a quantitative change of one attribute across matrix items and some items requiring deeper analytic reasoning, involving more than simple pairwise comparisons across figures (Panel C; see Carpenter et al., 1990). Across the 7 participants tested, both spatial and analytic items activated LPFC (Brodmann Areas 9, 46) over the control task (Panel A). Only the analytics activation was bilateral, and only it recruited additional lateral and inferior PFC regions (6, 10, 44, 45). Moreover, the analytic problems elicited greater activation than did the spatial problems in almost all of these PFC areas (6, 9, 44, 45, 46). Analytic problem solving, then, under the WM load of considering multiple relations among stimuli, broadly recruited LPFC into action.

The parametric manipulation of the relational complexity of RPM-like problems also selectively and lawfully increases LPFC activity. Christoff et al. (2001) presented 10 participants with items in which zero dimensions changed across stimuli (thus presenting identical items across the matrix), one dimension changed either horizontally or vertically, or two dimensions changed simultaneously, both horizontally and vertically. In a separate study, with eight participants, Kroger et al. (2002) varied the
Figure 9.1 Sample Modified Raven's Problems Used by Prabhakaran, Smith, Desmond, Glover, and Gabrieli (1997)

Note: (A) Control problems, in which participants selected one of eight choices as an identical match to the figure in the bottom right of the matrix. (B) Figural/spatial problems. (C) Analytical problems.
dimensional-change "load" of RPM-like items from zero to four (see Figure 9.2, Panels A and B, for examples). Behaviorally, load had big effects in both studies, with responses becoming slower and less accurate as relational complexity increased. In the imaging data, moreover, LPFC activation increased bilaterally with increasing complexity (especially in Areas 9, 10, and 46). As a separate manipulation of attention-control requirements, Kroger et al. also included test problems in which zero to four distractors were presented in the matrix (see Figure 2, Panel C). These distractors did not change within a matrix, and so they were clearly irrelevant to problem solution. Nonetheless, they parametrically impaired behavioral performance (i.e., more distractors were associated with larger impairments) and elicited
increasing bilateral activity in most of the same LPFC regions as did increasing complexity. What seems quite clear from these studies, then, as predicted by WM/attention theories of Gf and PFC functioning, is that LPFC regions are recruited in RPM-type problems as their memory load and control demands increase.

One of the three remaining studies examining nonverbal reasoning also adapted items from a common marker test of Gf, the Cattell Culture Fair test (Duncan et al. 2000). The test was a version of the odd-man-out task, in which four abstract figures/designs were presented in a row, and all but one was constructed according to the same abstract rule. The participant’s job was to find the odd figure. A control version of the task presented three identical figures/designs in each row, and so the odd figure was unmistakable. Although rCBF was measured only from the first 40 seconds of each 2-minute task, the reasoning task elicited greater LPFC activity than did the control tasks (bilaterally in Areas 8 and 46, left and right lateralized in 47 and 6, respectively).

Two additional spatial/figural studies have examined different forms of reasoning from those considered so far, but in my mind, they are sufficiently fluid and inferential in their requirements to include them here. Under PET, Wharton et al. (2000) presented 12 participants with a figural “analogies” task after about 40 minutes of task training; these participants were previously selected to be above average in IQ and above the college sophomore mean in verbal analogy performance. Each test item presented a source display with four shapes of different colors, locations, and textures, followed by a target display featuring four new shapes with either the same or different relations among them as the source (see Figure 9.3); participants made a yes/no decision about whether the target was analogous to the source. The control, “literal” condition presented either identical or nonidentical arrays on source and target displays. Analogy problems activated left LPFC (6, 10, 44, 45, 46, 47) relative to the control baseline.

In their study of transitive inference, Acuna, Eliassen, Donoghue, and Sanes (2002) also tested participants who previously demonstrated good performance in the critical task. Here, in behavioral sessions prior to scanning, 15 participants learned a serial sequence of 11 arbitrary shape/color combinations (e.g., red horizontal oval → light
blue vertical rectangle → orange circle → yellow square, etc.) through trial and error. In Phase 1 of training, each trial presented two colored shapes, taken from adjacent positions in the target sequence (e.g., orange circle/ yellow square), and participants guessed which was closer to the end of the sequence; accuracy feedback followed each response. After the learning criterion was reached, participants began the transitive inference task, in which shape pairs were presented that were separated by one, two, or three intervening shapes in the target sequence (e.g., red horizontal oval/yellow square), and so these pairs had never been tested during training; again, the participants had to choose the shape closer to the end of the sequence, but no feedback was provided. Participants had to infer the larger structure of the sequence from the individual, adjacent-paired comparisons they had made in Phase 1. Brain scanning under fMRI was then conducted on transitive inference problems with one intervening shape. Compared to a control task in which participants judged which of two shapes was taller than the other, the inference task evoked bilateral activation of LPFC (9, 46) and superior PFC (6, 8).

To summarize, fluid figural reasoning tasks, in which participants must infer a rule that simultaneously associates several novel objects, quite clearly and consistently activate LPFC structures, often bilaterally. Moreover, in cases where the memory and executive attention demands of these tasks vary parametrically, by increasing the number of object dimensions to be considered simultaneously or by presenting distractor objects, LPFC activity systematically increases with these demands.

Inductive and Deductive Verbal Reasoning

Vinod Goel and his colleagues have conducted much of the imaging research on complex verbal reasoning. Although these studies have not been designed to examine general fluid intelligence per se (the phrase is not mentioned in any of these publications), the novel reasoning involved, some of it inductive, suggests its relevance to this review. Interpretation of these studies must be a bit more provisional than that for the nonverbal studies considered above, however, because it has proven more difficult to establish an appropriate baseline against which to compare the reasoning task activation. All the control tasks have presented similar stimuli to those in the reasoning task, but the judgments required have been so different (e.g., counting the number of sentences that were about people) that one can never be sure exactly what has been subtracted out by the control task images. That said, the literature is somewhat consistent in indicating involvement of LPFC in a variety of verbal reasoning tasks. Where LPFC activity is seen, however, it is not always similarly lateralized across studies.

In formal deductive reasoning, participants must judge whether a conclusion follows logically from its premises. Most deductive reasoning studies have used three-term problems, with two premises followed by a valid or invalid conclusion, and most have found LPFC activity. As the first demonstrations, Goel, Gold, Kapur, and Houle (1997, 1998) tested 10 and 12 participants, respectively, who were selected for prior good performance in syllogistic reasoning (e.g., All men have sisters. Spearman was a man. ∴ Spearman had a sister). Under PET, reasoning activated left LPFC areas (45, 46, 47) over the “people judgment” baseline.

Subsequent to these initial studies, Goel and colleagues have used fMRI techniques to explore a range of content variables within deductive reasoning, and they have been rather inconsistent in the laterality of the PFC activation elicited. For example, Goel and Dolan (2000) presented 10 participants with rows of three imaginary animal-like stimuli, and participants were told that if, and only if, the animals all had the same abdomen and tail, then they were the same kind of animal. Such same/different judgments activated LPFC bilaterally (9, 46) and in the right hemisphere (44), compared to a baseline task of judging whether the three animals were presented in a straight horizontal line.

Using more traditional deductive tasks, Goel, Buchel, Frith, and Dolan (2000) and Goel and Dolan (2001) had 11 and 14 participants, respectively, judge the validity of syllogisms with and without semantic content (no-content problems took the form of All Ps are Bs, etc.). In both studies, the baseline condition presented conclusions that were completely unrelated to
the premises, and so participants could immediately respond to them in the negative. Goel et al. (2000) found content problems to activate left LPFC (44) and no-content problems to activate LPFC bilaterally (44, 45). Similarly, Goel and Dolan (2001) found left PFC (9) activation for reasoning collapsed over content and no-content problems (along with bilateral medial PFC in Area 6), but neither content nor no-content problems alone activated LPFC over baseline. So, there seem to be some inconsistencies in the laterality of activation, as well as the particular Brodmann areas involved, across these deductive reasoning studies. Although it seems as though increasing verbal/semantic content increases left laterality, the database is currently too small for much confidence in any such speculation.

Exacerbating the confusion, a deductive reasoning PET study using only slightly different stimuli activated right LPFC (9, 44) over baseline. Parsons and Osherson (2001) scanned 10 participants while they solved syllogisms using propositional stimuli (e.g., Either he likes country music or listens to opera. He does not like country music. ∴ He listens to opera), which should minimize spatial, Venn diagram–type solution strategies. As such, the right hemisphere activation they found is even more puzzling. It may be that participants still were able to use spatial strategies here, or it may be that the baseline condition used was somehow responsible (participants detected whether any of the statements had anomalous content, such as: If he is divorced then he has never been married). Indeed, Osherson et al. (1998) found syllogistic reasoning to not activate LPFC at all against a similar anomaly judgment baseline.

One suggestion that the anomaly baseline task might be the culprit here is that Goel et al. (2000) and Goel and Dolan (2003) found that problems with “incongruent,” anomalous content activated right LPFC (46/45) relative to those with “congruent” or sensible content. More work will obviously be required to sort through such laterality and baseline issues. For now, however, the evidence suggests some uniformity in LPFC activity during deductive reasoning, but not nearly as much as in the figural inductive reasoning literature reviewed previously.

Studies of inductive and relational reasoning using verbal propositions have yielded clearer and more consistent results regarding PFC activation. Goel et al. (1997) found that when participants were asked to judge whether a conclusion was plausible given two premises, the lateral, superior, and inferior areas of PFC were activated in the left hemisphere (8, 9, 10, 47). Similarly, Parsons and Osherson (2001) asked participants whether a conclusion following two premises was more likely to be true than false, and they also found left lateral and inferior PFC activity (9, 44). Results from relational reasoning studies converge with those above. Goel et al. (1998) had participants judge the validity of conclusions about the spatial arrangement of subjects (e.g., Officers stand next to generals, etc.) or other, nonspatial relations (e.g., Officers are younger than generals, etc.) and found both to activate left lateral and inferior PFC regions (45, 46). Finally, Ruff, Knauff, Fangmeier, and Spree (2003) auditorially presented their 12 participants with propositions regarding the spatial layout of colored geometrical shapes and found bilateral activation of LPFC (Area 10) versus a resting baseline. The extra, right hemisphere activation here may have resulted from the spatial thinking required in the absence of other semantic content in the stimulus materials.

In summary, then, studies of the brain mechanisms engaged in deductive and inductive verbal reasoning largely elicit activity in LPFC areas, concentrated chiefly in the left hemisphere, as expected in language tasks. At least we can say this with some confidence regarding inductive and relational reasoning; the inductive relational data are internally consistent in the verbal domain, and they also corroborate the data from the nonverbal, figural domain reviewed previously (with the exception of generally greater bilateral activity in the nonverbal domain). Although more work is required to elucidate the prefrontal neural mechanisms involved in formal deduction, it is clear that inductive fluid reasoning tasks selectively and consistently engage cells in LPFC.

Prefrontal Contributions to Individual Differences in Reasoning

As I argued in introducing this literature review, knowing that certain brain areas contribute
to the performance of reasoning tasks does not necessarily indicate that these same areas are critical to \textit{individual differences} in reasoning within a task or across many tasks. Thus, to more fully inform theories of Gf, neuroimaging work must test for the relation between brain activity and performance within a task, as well as contrasting neural activity and performance across tasks.

\textit{Prefrontal Activation and Individual Differences Within Tasks}

Very few studies have looked for an association between the strength or extent of neural activation in a brain region and performance on a reasoning task, and the results have been mixed. Risberg and Ingvar (1973), for example, found RSPM performance to be unrelated to rCBF in any individual brain area, whereas Gur et al. (1994) found verbal analogies performance to correlate with rCBF in all measured brain areas ($r = .40$). Nonetheless, in two PET studies, we can find some evidence for a special role for LPFC in Gf differences. Haier et al. (1988) seem to have found, despite the poor spatial resolution of their data, that absolute metabolic rates in lateral/superior PFC (as well as in parietotemporal and occipital cortices) correlate negatively with RAPM scores ($r = -.70$ to -.80), such that poorer performers showed greater activation of these areas overall. However, relative metabolic rates in lateral/superior PFC were unique in showing a substantial positive association with performance ($r = .40$ to .60), indicating that participants who recruited LPFC areas to a greater degree than predicted by their whole-brain activity were the better performers.

In the lone cognitive aging study in this literature, Esposti et al. (1999) tested 41 participants between 18 and 80 years old in the RSPM and found that age correlated negatively with performance but positively with task-related rCBF in LPFC. That is, older participants performed worse on the RSPM despite engaging LPFC to a greater degree. This finding, along with the Haier et al. (1988) negative correlations between RAPM and absolute PFC metabolism, suggests that PFC is recruited to solve inductive reasoning problems under worst-case conditions, such as when problems are most difficult or when one has reduced fluid abilities.

\textit{Prefrontal Activation and Individual Differences Across Tasks}

A multivariate approach to neuroimaging Gf examines the relation between regional brain activity and performance across multiple tasks. Because Gf is a multivariate construct, such studies should best inform the effort to link Gf with one or more neurological substrates. Unfortunately, only three such multivariate studies appear to have been conducted to date. One is inconclusive for present purposes because the variables studied were not linked to PFC activation (Haier, Siegel, Tang, Abel, & Buchsbaum, 1992), and another demonstrated brainwide correlations not specific to PFC (Larson, Haier, LaCasse, & Hazen, 1995). In the latter study, high and average Raven’s scorers ($M_s = 30.4$ and 22.0, respectively, out of 36) completed both a hard version of the backward digit span task (set sizes tailored to yield 75% accuracy) and an easy version of the task (tailored to yield 90% accuracy) under PET. In both hard and easy tasks, high Raven’s scorers showed greater rCBF in all brain lobes than did low scorers (discriminations smaller than entire lobes were not made); low scorers were also characterized by a smaller global increase in rCBF for hard versus easy problems.

The good news for this literature is that one multivariate study, using state-of-the-art fMRI methodology, has very recently provided evidence for the importance of LPFC to individual differences in Gf. It provides a model for future research in this area and also nicely ties together the theoretical ideas about WM capacity, executive attention, Gf, and PFC functioning that I described at the outset of this chapter.

Gray, Chabris, and Braver (2003) tested 48 participants in two cognitive tasks reflecting Gf (RAPM) and WM ($n$-back), respectively, and assessed the relation between RAPM performance (outside the scanner) and individual differences in $n$-back performance and neural activity. In the $n$-back, participants viewed a sequence of words or faces at a 2.36-second rate in
the scanner and indicated whether each stimulus matched the one seen three items ago. Recall from my discussion of the PFC/WM literature that the n-back task is considered a marker of WM because it requires participants to continually maintain and update an evolving list of stimuli. A large neuroimaging literature indicates that as the value of n increases between 1 and 3, so does LPFC activity.

As illustrated in Figure 9.4, Gray et al. (2003) further assessed the extent of executive attention in the n-back task by contrasting two different kinds of nontargets—"lure" and "nonlure" foils. Lure foils matched a stimulus two, four, or five items back, and so they elicited conflict with the three-back demand. Nonlure foils were novel, one-back matches, or ≥ six-back matches and so provoked less interference (indeed, behavioral measures indicated that lure foils were less accurately rejected than were nonlure foils). Behaviorally, lure foil accuracy correlated positively with Gf (r = .36), even after partialling out either nonlure foil or target accuracy (.27 and .26, respectively). Gf clearly predicted executive attention effectiveness in the n-back task, as indexed by interference resistance.

Gray et al. (2003) then examined the event-related brain activity during n-back performance that was specific to resolving lure foil interference. They did so, first, in brain regions expected a priori to be important to conflict resolution—LPFC, anterior cingulate, and lateral posterior cerebellum—and, second, via a brainwide search for event-related n-back activity. In the a priori search, lure trial activity in LPFC correlated significantly with both lure trial accuracy and Gf, as did activity in some inferior frontal areas and anterior cingulate (rs = .40-.50). The whole-brain analysis demonstrated similarly strong correlations between Gf and lure-related activity in regions of frontal, parietal, and temporal cortices. In fact, lure trial neural activity in LPFC and the parietal cortex together accounted for 99.9% of the variance shared by Gf and lure trial accuracy! Figure 9.5 illustrates the relationship between lure accuracy and Gf within left LPFC. High-Gf participants showed a greater increase in activation on lure trials than did low-Gf participants. In fact, many low-Gf participants showed a decrease in activation on lure trials relative to other trials. Similar to Kane and Engle's (2000) finding regarding WM differences and PI susceptibility, high-Gf individuals engaged in more extensive "mental work" in the presence of interference than did those of low Gf.

Gray et al. did not find strong correlations between RAPM and brain activity for other aspects of n-back performance. Although RAPM scores correlated with n-back accuracy on nonlure and target trials, this correlation with brain activity was much weaker. Despite the fact that LPFC and parietal areas generally increased their activity in the n-back task compared to baseline, none of this sustained brain activity correlated significantly with Gf, thus recalling the analogy of distinguishing basic braking
ability. Consistent with some work from the neuropsychological tradition, in which behavioral deficits are linked with particular patterns of brain damage. LPFC areas appear to be particularly important to the performance of Gf tasks. In both verbal and nonverbal domains, inductive, relational, and transitive (and sometimes deductive) reasoning problems elicit LPFC activity. Of special importance, no other brain regions have been identified in the literature that are so consistently involved across the variety of reasoning and stimulus domains reviewed here.

This research area is growing and producing interesting and replicable results. However, from the perspective of trying to inform our understanding of intelligence, it is time to move beyond examining brain activity during individual reasoning tasks. The multivariate nature of the Gf construct demands that some connection be made between activity and performance in one task and activity and performance in others. A recent study by Gray et al. (2003) moves the field forward by linking reasoning performance and brain activity across multiple tasks and by using marker tasks of theoretically important constructs, such as WM and attention control. Their important findings support the behavioral associations among WM, attention, and Gf constructs and suggest that PFC activity related to

Figure 9.5 Neuroimaging Data From Gray, Chabris, and Braver (2003)

Note: (A) Correlation between Raven’s Advanced Progressive Matrices (APM) score and functional magnetic resonance imaging (fMRI) signal from lateral prefrontal cortex (LPFC) during correctly rejected three-back lure trials. (B) Time course of fMRI signal from the LPFC during correctly rejected three-back lure trials for high- versus low-Gf participants.

mechanisms from sources of braking variation. In sum, only WM processes that were tied directly to interference control were supported by Gf-related brain activity.

CONCLUSION

Neuroimaging research has begun to reveal the brain structures important to general fluid
attention control is what binds these constructs together. Future work that continues to combine cognitive-behavioral, psychometric, and neuroscience theory and technique holds great promise in advancing our understanding of not only how minds work but also how they differ from one another and why.

NOTES

1. In parallel with the figural task, Duncan et al. (2000) also tested participants with an odd-man-out modification of the Educational Testing Service (ETS) letter series test (Ekstrom, French, Harman, & Dermen, 1976). Here, four sets of letter sequences were presented in a row, and all but one was constructed according to the same abstract rule. Again, a control version of the task served as a baseline, in which the to-be-discovered rule was relatively obvious (three sequences were alphabetical). In contrast to the figural task, the letter task yielded only left lateralized activity (10/46, 46, 47). Although the verbal nature of the task may account for its laterализation, it was also true that the verbal reasoning and control tasks had more similar g factor loadings to each other than did the spatial tasks (in a pilot study), and so the control condition in the letter task may have been more demanding and a less appropriate baseline. What is clear, however, is that both the inductive reasoning tasks evoked lateral prefrontal cortex (LPFC) activity similar to that seen in the Raven's Progressive Matrices (RPM) test.

2. The first imaging study of verbal reasoning that I am aware of examined 34 adults in an analogies test (Gur et al., 1994). The spatial resolution for the regional cerebral blood flow (rCBF) measurement was poor, but one detector appears to have been placed over LPFC Area 9; in fact, this area, along with other prefrontal cortex (PFC) and posterior structures, showed elevated activity during the task compared to a resting baseline.

3. Esposito, Kirkby, Van Horn, Ellmore, and Berman (1999) did not report a relative measure of regional rCBF like that from Haier et al. (1988), and so it is unclear whether that finding is also replicable.

4. Haier, Siegel, Tang, Abel, and Buchsbaum (1992) tested participants in the Raven's Advanced Progressive Matrices test in a session outside the scanner, and these participants were given extensive practice on a computerized visuospatial-motor game called Tetris. Although the correlation between Raven's score and whole-brain metabolism during Tetris performance showed the predicted reduction following practice (naive r = .77, practiced r = .09), Raven's scores continued to predict Tetris performance (naive r = .36, practiced r = .41), and LPFC activity during Tetris did not change with practice, remaining substantial at both time periods. No correlations were reported between Raven's performance and regional brain activity in either naive or practiced performers; only naive-to-practiced change measures were reported.

REFERENCES


by neurons in the primate prefrontal cortex. *Nature*, 393, 577–579.


