

A Neurocomputational Model of Analogical Reasoning and its Breakdown in Frontotemporal Lobar Degeneration

Robert G. Morrison¹, Daniel C. Krawczyk¹, Keith J. Holyoak¹,
John E. Hummel¹, Tiffany W. Chow², Bruce L. Miller³,
and Barbara J. Knowlton¹

Abstract

■ Analogy is important for learning and discovery and is considered a core component of intelligence. We present a computational account of analogical reasoning that is compatible with data we have collected from patients with cortical degeneration of either their frontal or anterior temporal cortices due to frontotemporal lobar degeneration (FTLD). These two patient groups showed different deficits in picture

and verbal analogies: frontal lobe FTLD patients tended to make errors due to impairments in working memory and inhibitory abilities, whereas temporal lobe FTLD patients tended to make errors due to semantic memory loss. Using the “Learning and Inference with Schemas and Analogies” model, we provide a specific account of how such deficits may arise within neural networks supporting analogical problem solving. ■

INTRODUCTION

Analogy is ubiquitous in human learning and discovery (Gentner, Holyoak & Kokinov, 2001; Holyoak & Thagard, 1995), and has long been viewed as a core component of intelligence (Sternberg, 1977; Spearman, 1923). Research on analogies in the psychometric tradition has focused on four-term problems, for example:

BLACK:WHITE::NOISY:?(1) QUIET(2) NOISIER.

Several processes are required to solve such problems successfully. The reasoner needs to: (1) encode the terms of the analogy; (2) search semantic memory for a relation that connects the first two terms, BLACK and WHITE (opposite); (3) bind this relation to the specific concepts being related in working memory (i.e., form the proposition “BLACK is the opposite of WHITE”); and (4) map the resulting proposition to the third and possible fourth terms to infer an answer (BLACK maps to NOISY and WHITE maps to QUIET, based on the proposition “NOISY is the opposite of QUIET”). Analogical reasoning thus requires the ability to retrieve information from semantic memory, together with the ability to form and manipulate mental representations of relations between objects and events in working memory.

The goal of the present article is to present evidence concerning the breakdown of the component processes of analogical reasoning in patients suffering from cortical degeneration, and to simulate the findings using a

neurocomputational model of analogical reasoning, “Learning and Inference with Schemas and Analogies” (LISA; Hummel & Holyoak, 1997, 2003). Previous research has shed some light on the neural substrate of analogical reasoning, and numerous computational models of analogy have been developed. However, the links between detailed process models of high-level reasoning and investigations of its neural basis have been tenuous. Our aim is to provide an example of how a computational model of reasoning can be used to help understand the processes that operate at the neural level to enable complex reasoning in normal individuals, and how frontal versus temporal lobe degeneration may selectively impair these processes.

Neural Substrate for Analogical Reasoning

Both computational analyses and empirical evidence suggest that analogical reasoning depends on working memory. For example, experiments utilizing dual-task methodology have shown that the processes of binding and mapping used in analogical reasoning require working memory (Morrison, Holyoak, & Truong, 2001; Waltz, Lau, Grewal, & Holyoak, 2000; Baddeley, Emslie, Kolodny, & Duncan, 1998). Other studies have shown that working memory is at least in part realized in subareas of the prefrontal cortex, which have been implicated both in tasks that involve simple short-term maintenance and in more complex manipulation tasks characteristic of the central executive (see Kane & Engle, 2003; Miller & Cohen, 2001; D’Esposito, Postle, & Rypma, 2000, for reviews).

¹University of California, Los Angeles, ²University of Toronto, ³University of California, San Francisco

Recent studies have implicated the prefrontal cortex specifically in reasoning tasks. Wharton et al. (2000) used positron emission tomography to investigate brain activation associated with solving geometric analogy problems. Participants judged whether two sets of objects were identical (literal condition) or analogous (analogy condition). Regions active during analogy but not during literal comparison included the left dorsomedial prefrontal cortex, Brodmann's area (BA) 44 and 45 (also see Luo et al., in press). Other studies have examined the Ravens Progressive Matrices test (RPM; Raven, 1938), a task cognitively and psychometrically similar to analogical reasoning (Snow, Kyllonen, & Marshalek, 1984). Waltz et al. (1999) found that frontal lobe FTLD patients showed a marked deficit in RPM-type problems that required integration of two relations compared with normal controls and FTLD patients with anterior temporal lobe damage. The frontal lobe FTLD patients performed comparably to the other groups on less complex RPM-type problems that could be solved using zero or one relation. Waltz et al. interpreted this result, as well as similar findings with deductive reasoning problems, as evidence that the prefrontal cortex is important for the integration of multiple relations, a process required to solve reasoning problems that are relatively complex (Robin & Holyoak, 1995; Halford, 1993; Halford, Wilson & Phillips, 1998). Several studies using functional magnetic resonance imaging (fMRI) have shown bilateral prefrontal cortex activation during RPM problem solving (Prabhakaran, Smith, Desmond, & Glover, 1997) with progressively greater involvement of the right dorsolateral and anterior prefrontal cortex at higher levels of relational complexity (Kroger et al., 2002; Christoff et al., 2001). Integrating relations requires the brain to rapidly acquire new connections (Hummel & Holyoak, 1997, 2003), a capability that has been described as one of the core functions of the prefrontal cortex (Miller & Cohen, 2001; Rolls, 1996). Cell recordings with primates (Asaad, Rainer, & Miller, 1998), as well as human studies (Cools, Clark, Owen, & Robbins, 2002; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001), have demonstrated the underlying neural correlates of such rapid learning associated with the prefrontal cortex.

Processing analogies also is likely to utilize inhibitory mechanisms to manage information in working memory. Considering our earlier verbal analogy example:

BLACK:WHITE::NOISY:?(1) QUIET(2) NOISIER,

if the semantic association between NOISY and NOISIER is stronger than that between NOISY and QUIET, the correct analogical response, then QUIET may initially be less active than the distractor item, NOISIER. Thus, during analogical reasoning, it may be important to inhibit information that is semantically related to the analogy but competes against the relational match. Behavioral studies of analogical mapping have shown that the ability to identify relational correspondences is

impaired when competing featural matches are present (e.g., Holyoak & Koh, 1987; Ross, 1987, 1989; Gentner & Toupin, 1986). More generally, many researchers have suggested that inhibition is an important mechanism for complex cognition (see Kane & Engle, 2003; Dempster & Brainerd, 1995; Dagenbach & Carr, 1994, for reviews), and that changes in inhibitory control may explain important developmental trends (Bjorklund & Harnishfeger, 1990; Diamond, 1990; Hasher & Zacks, 1988) and individual differences (Kane & Engle, 2003; Dempster, 1991) in complex cognition.

Many researchers have viewed inhibition as an important function of the prefrontal cortex (see Kane & Engle, 2003; Miller & Cohen, 2001; Shimamura, 2000, for reviews). Many complex executive tasks associated with frontal lobe functioning (e.g., Tower of Hanoi or London, Wisconsin Card Sorting) have important inhibitory components (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Welsh, Satterlee-Cartmell, & Stine, 1999). Shimamura (2000) has suggested that the role of the prefrontal cortex is to dynamically filter information, a process that requires the use of both activation and inhibition to keep information in working memory relevant to the current goal. Miller and Cohen (2001) have suggested that "the ability to select a weaker, task-relevant response (or source of information) in the face of competition from an otherwise stronger, but task-irrelevant one [is one of the most] fundamental aspects of cognitive control and goal-directed behavior" (p. 170), and is a property of the prefrontal cortex.

An addition to working memory and inhibitory mechanisms, analogical reasoning also requires the ability to access semantic information in order to retrieve and encode the relations relevant to the analogy. While there is considerable debate about how semantic information is stored in the brain (Markowitsch, 2000), it seems clear that the temporal lobes are critical for the preservation of semantic memory (Hodges, 2000) and the relational information necessary for analogical reasoning. The anterior temporal cortex appears to be particularly important for verbally mediated conceptual knowledge (Mummery et al., 1999; Martin, Wiggs, Ungerleider, & Haxby, 1996).

Overview of the Present Study

In this study, we report two experiments investigating the role of the prefrontal cortex and the anterior temporal cortex in analogical reasoning, and a LISA simulation of the results of the second study. These experiments were performed with patients diagnosed with frontotemporal lobar degeneration (FTLD). FTLD is a regional neurodegenerative etiology of dementia distinct from Alzheimer's disease (e.g., Hodges & Miller, 2001; Neary et al., 1998). Based on clinical features and neuroimaging (see Figure 1), FTLD patients can be divided based on their primary locus of damage. Frontal lobe FTLD patients

typically present with behavioral disturbances and have damage to the orbital medial regions of the prefrontal cortex, which extends into the dorsolateral prefrontal cortex and inferior frontal gyrus as the disease progresses (Hodges & Miller, 2001). Temporal lobe FTLTD presents with a loss of conceptual knowledge, fluent aphasia, and late emotional changes. Damage is most apparent in the anterior temporal cortex (Hodges & Miller, 2001). Studies of FTLTD patients can provide a useful complement to neuroimaging studies by making it possible to investigate the global roles of frontal and anterior temporal lobes in reasoning. The present study compares the performance of frontal and temporal lobe FTLTD patients to that of age- and education-matched controls.

Analogy Experiments

Both experiments involved solving analogy problems. In Experiment 1, participants answered analogy problems based on a paradigm initially developed by Markman

and Gentner (1993). On each trial in this task, participants first studied pairs of pictures containing multiple objects and at least one dominant relation, which is present in both pictures (see Figure 2 for an example). The experimenter then indicated one object in the top picture and asked the participant to select one object in the bottom picture that “goes with” the indicated object. In each case, the indicated object could plausibly be mapped to two different objects in the bottom picture—one that matches based on featural attributes (*a featural match*), and one that matches based on its relation(s) to other object(s), that is, based on an analogy between the two pictures (*a relational match*). Typically, participants select either the featural or relational match. To identify the relational match for these complex visual analogies, it will generally be necessary to encode and manipulate multiple relations among the depicted objects. Using this task, Waltz et al. (2000) demonstrated that dual tasks designed to interfere with working memory decrease the number of relational

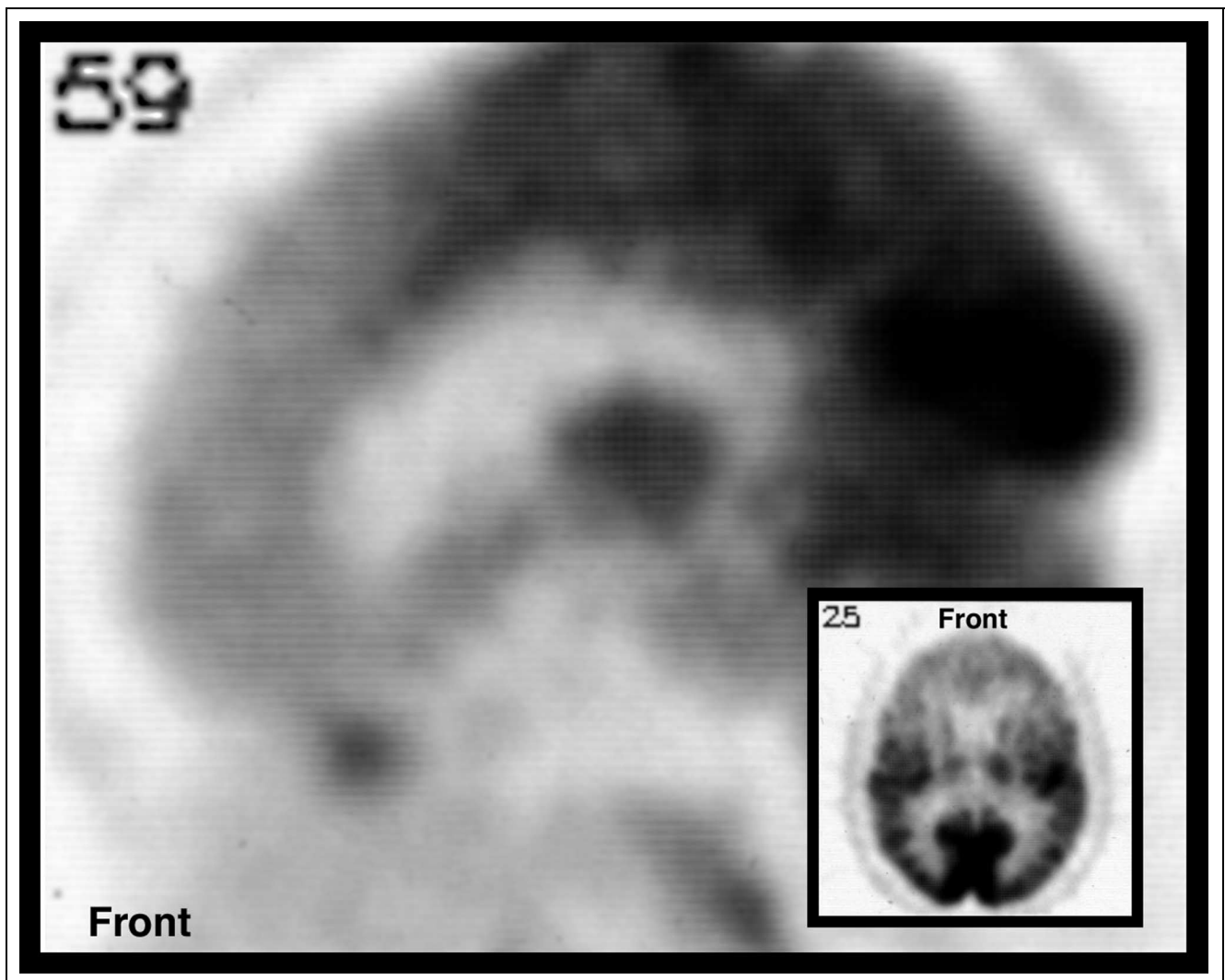


Figure 1. Positron emission tomography scan from patient HS at rest. HS, who was diagnosed with frontal lobe FTLTD (later autopsy confirmed), participated in both Experiments 1 and 2.

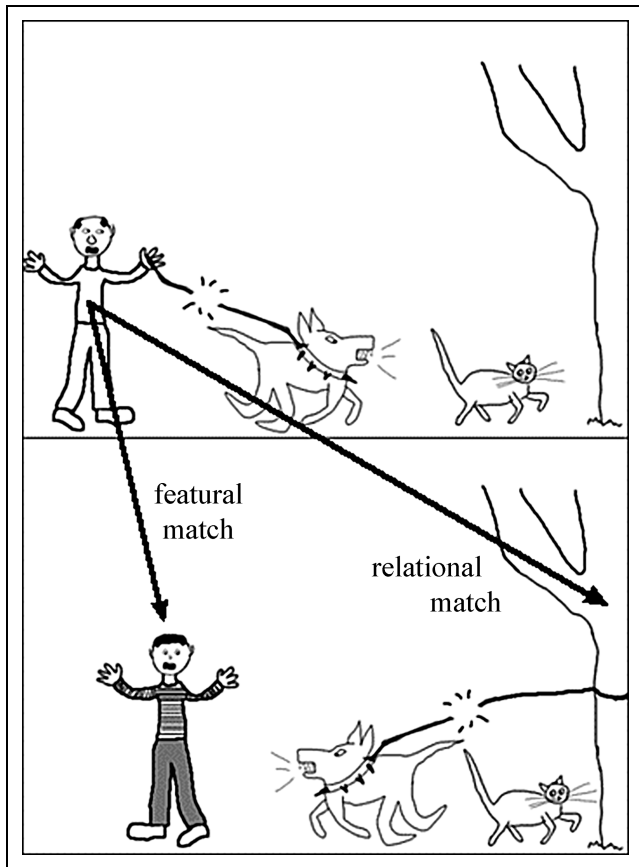


Figure 2. Example problem from picture analogy task (Experiment 1). Adapted from Tohill and Holyoak (2000).

responses given, with a corresponding increase in the number of featural responses. In addition to the working memory resources needed to represent and manipulate relations, this task likely requires inhibition of the salient featural match in order to facilitate relational processing. We hypothesized that frontal lobe FTLD patients will give fewer relational responses because they have diminished ability to represent relations and learn analogical correspondences, and because they will have difficulty inhibiting the salient featural match in order to produce a less-salient relational response.

In addition, the picture analogy task also requires retrieval of conceptual knowledge from semantic memory to encode the relations necessary to process the analogy. We hypothesized that temporal lobe FTLD patients would favor featural over relational responses due to the perceptual similarity of featural matches to the indicated objects and the patients' diminished semantic memory, making access to the relational information necessary to process the analogy more difficult.

In Experiment 2, we examined a verbal analogy task that allowed us to better characterize the specific analogy processing failures in the frontal and temporal lobe FTLD patients. Problems of the form A:B:C:D or D' (adapted from Sternberg & Nigro, 1980) were used. These prob-

lems varied the association between C:D and C:D' word pairs. A semantic facilitation index (SFI) was calculated for each problem to characterize the association of the correct relational pair (C:D) relative to the distractor pair (C:D'). For example, for the problem

PLAY:GAME::GIVE:?(1) PARTY(2) TAKE

the C:D pair (GIVE:PARTY) is less associated than is the C:D' pair (GIVE:TAKE), giving the problem a negative SFI score. The problems were divided into three groups: negative SFI, neutral SFI, and positive SFI.

We predicted a different pattern of results for frontal and temporal patients. Because these simple four-term analogy problems are each based on a single relation between the A and B terms (unlike the more complex picture analogies used in Experiment 1), it should be possible for frontal patients to perform the basic analogical mapping despite their diminished working memory. In the positive and neutral SFI conditions, the analogical answer (D) does not face competition from a competitor (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 generate the analogical answer, it will be necessary not just to activate it, but to rapidly learn the analogical correspondences so that activation of the analogical answer D can be maintained in the face of automatic activation of the foil D'. It will then be necessary to inhibit the D' response so that the analogical response D can be made. Because of their postulated deficits in both rapid learning and inhibitory control, we predicted that frontal patients would be selectively impaired in the negative SFI condition relative to the positive and neutral SFI conditions. In contrast, we predicted that temporal patients would show a more uniform decline in verbal analogy performance across all three conditions due to their loss of the conceptual information necessary to encode the relations in the analogy problem.

LISA: A Neurocomputational Model of Analogical Reasoning

In an effort to substantiate the importance of semantic memory, working memory, and inhibition in analogical reasoning, we modeled the deficits in patient performance in LISA (Hummel & Holyoak, 1997, 2003), a symbolic-connectionist model of analogical reasoning. Although numerous computational models of analogy have been developed (see French, 2002, for a review), no other such model has been used to simulate neuropsychological data.

LISA represents propositions using a hierarchy of distributed and localist units (see *Methods*, Figure 5). LISA includes both a long-term memory for propositions

and concept meanings and a limited-capacity working memory. LISA's working memory representation, which uses neural synchrony to encode role-filler bindings, provides a natural account of the capacity limits of working memory because it is only possible to have a finite number of bindings simultaneously active and mutually *out* of synchrony (see Appendix A in Hummel & Holyoak, 2003, for details). 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, remaining in active memory when they are not currently being fired. LISA also includes a set of *mapping connections* between units of the same type (e.g., object, predicate) in separate analogs. These connections grow whenever the corresponding units are active simultaneously, and thereby permit LISA to learn the correspondences between structures in separate analogs. They also permit correspondences learned early in mapping to influence the correspondences learned later. Hummel and Holyoak (1997, 2003) hypothesized that the rapid learning of mapping connections is an important function of working memory as implemented in the prefrontal cortex, consistent with recent work by Asaad et al. (1998). The behavioral experiments and computational simulations reported here serve as additional tests of this hypothesis.

In addition to the mapping connections, several other aspects of LISA's operation also map onto functions associated with the prefrontal cortex and anterior temporal cortex. Inhibition (considered an important function of the prefrontal cortex) plays a central role in several aspects of LISA. These include: (a) LISA's ability to select items for placement into working memory; (b) its working memory capacity for role-filler bindings; (c) 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); and (d) 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 (see Holyoak & Thagard, 1989). Damage to the prefrontal cortex is expected to adversely affect all these aspects of LISA's ability to solve analogies. Damage to the anterior temporal cortex—and the resulting loss of semantic knowledge—corresponds in LISA to a loss of connections between the semantic properties of relational roles (i.e., semantic units) and units representing those roles in long-term memory.

RESULTS

Experiment 1: Picture Analogy Task

The results from the picture analogy task are summarized in Figure 3. Frontal lobe FTLD patients chose fewer relational responses and more featural responses than control participants, $t(18) = 5.2, p < .001$ and $t(18) = 4.0, p < .001$, respectively. Frontal lobe FTLD patients showed a trend toward choosing more unrelated choices than control participants, but this trend was not reliable, $t(16) = 1.6, p = .14$. Likewise, temporal lobe FTLD patients also chose fewer relational responses and more featural responses than did control participants, $t(18) = 3.8, p = .001$ and $t(18) = 3.0, p = .007$, respectively. Temporal lobe FTLD patients chose no more unrelated objects than did control participants, $t(18) = -0.527, ns$.

When frontal lobe FTLD patients did report relational answers, they usually were solving problems in which the relational match was also spatially aligned (e.g., objects on the left in both top and bottom pictures), or in which the target object and the relational match

Figure 3. Results of picture analogy task (Experiment 1). Both frontal and temporal lobe FTLD patients show a reliable decrease in the number of relational responses, with a corresponding increase in featural responses, compared to age- and education-matched control participants.

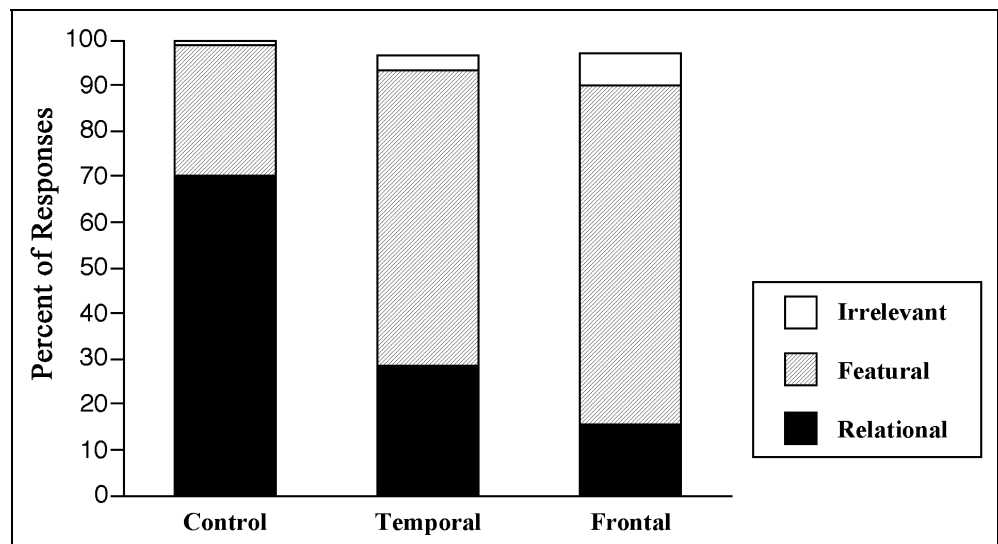
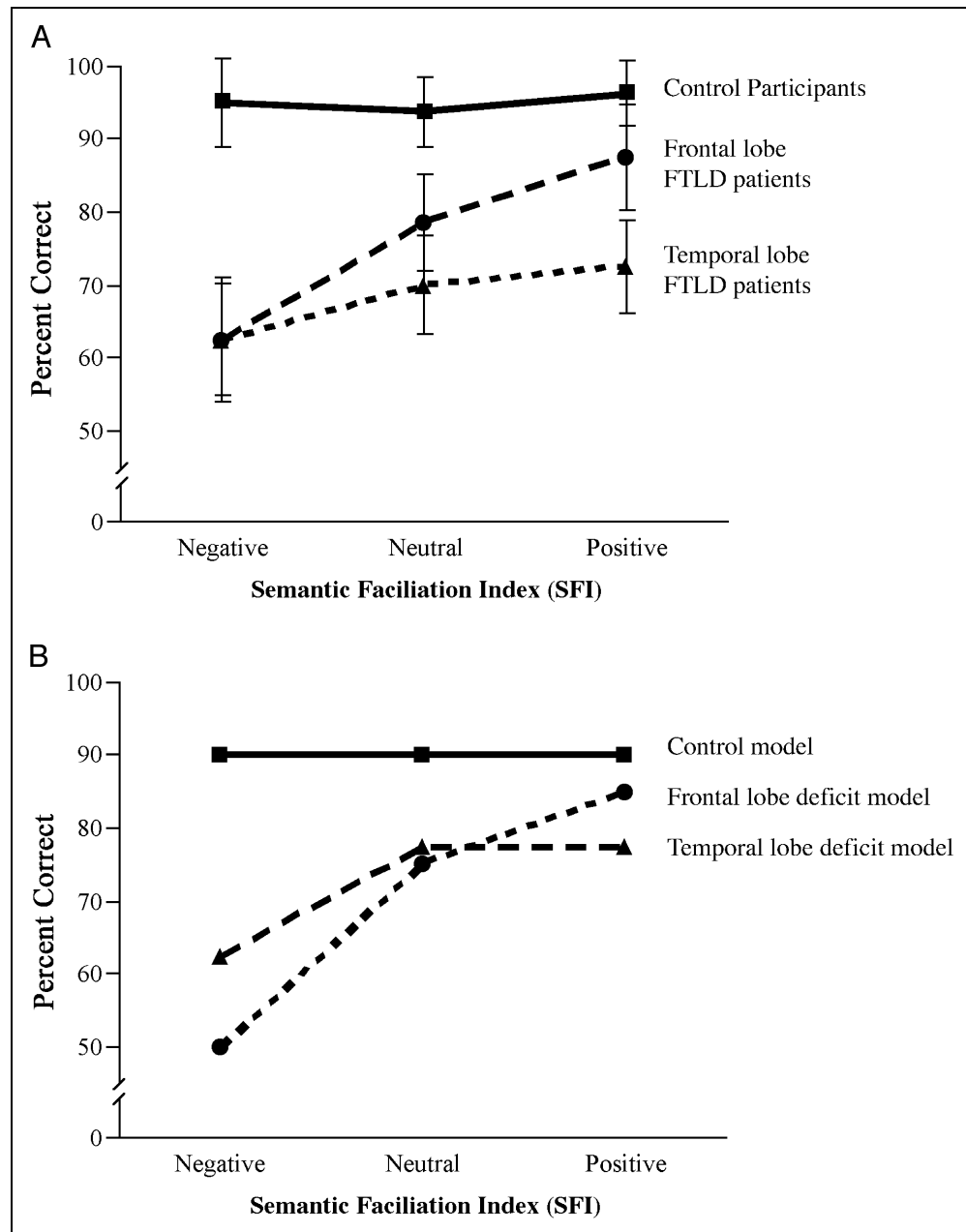


Figure 4. Results of verbal analogy task (Experiment 2). (A) Temporal lobe FTLD patients show a uniform decrease in verbal analogy performance regardless of SFI problem type, whereas frontal lobe FTLD patients show a greater deficit in negative SFI problems than in neutral or positive problems. Error bars reflect ± 1 standard error of the mean. (B) Simulation of patient performance using the LISA model.



also shared some perceptual similarity. This finding suggests that frontal lobe FTLD patients were more likely to choose the relational answer when it was supported by perceptual cues, minimizing the need to inhibit competing information. Temporal lobe FTLD patients distributed their relational responses more evenly across the set of 10 picture pairs.

Experiment 2: Verbal Analogy Task

The results from the verbal analogy task are summarized in Figure 4A. Frontal lobe FTLD patients made more errors than did control participants, $F(1,19) = 11.5$, $MSE = 0.038$, $p = .003$. Likewise, temporal lobe FTLD patients also made more errors than did control partic-

ipants, $F(1,19) = 18.8$, $MSE = 0.038$, $p < .001$. Testing our specific hypothesis, we found that frontal lobe FTLD patients showed a specific deficit in negative SFI problems relative to neutral and positive SFI problems relative to control participants, $F(1,19) = 5.0$, $MSE = 0.140$, $p = .04$. In contrast, temporal lobe FTLD patients did not show this pattern, $F(1,19) < 1$, $MSE = 0.140$, ns .

Simulation of Verbal Analogy Task in LISA

LISA simulations were performed for the verbal analogy task, which allowed better control of stimulus variables than did the picture analogy task. Our intent was to model three key characteristics of the verbal analogy data presented in Figure 4A: (1) control participants

show no effect of SFI problem type; (2) temporal lobe FTLN patients show depressed performance that is relatively unaffected by SFI problem type; and (3) frontal lobe FTLN patients show a specific deficit on negative SFI problems relative to neutral and positive SFI problems. A full description of the assumptions made in modeling the verbal analogy task is presented in the *Methods* section of this article. In brief, we created representations for the various SFI problem types by assuming word pairs that are highly associated in long-term memory have more relational instances (i.e., propositions explicitly relating them) in long-term memory relative to less associated word pairs. Frontal lobe FTLN patient performance was simulated in LISA by decreasing the learning rate for mapping connections and recipient inhibition level. In contrast, temporal lobe FTLN patient performance was simulated by increasing the semantic death rate (the probability, p , of losing a connection between a semantic unit representing a relational role and a predicate unit for that role). These alterations were based on the current conceptions of the deficits caused by these two subtypes of FTLN: Frontal lobe damage results in the loss of ability to rapidly form associations and inhibit irrelevant stimuli, whereas temporal lobe damage leads to semantic memory loss.

Because of the stochastic nature of the simulations, we ran 40 trials per problem type with model parameter settings corresponding to the various patient groups. LISA's mean percent correct for each of the conditions is shown in Figure 4B; the model's performance can be compared with the patient data shown in Figure 4A. The simulation results, like the patient data, show all three patterns described above. Additional simulations revealed that to simulate the performance of frontal lobe FTLN patients, it is in fact necessary to impair *both* LISA's rapid learning and its recipient inhibition; neither factor alone is sufficient.

DISCUSSION

Our goal in this study was to investigate the neural basis of analogical reasoning. In two experiments, we found that frontal and temporal lobe FTLN patients both exhibit deficits in analogical reasoning relative to age-matched controls, but that the underlying basis for their deficits appears to differ. When given pairs of complex pictures (Experiment 1), both patient groups were impaired in selecting a relation-based match for an object in the presence of a salient feature-based match. Although their diminished level of relational responses was comparable, there was some evidence that the underlying deficit differed for the two groups. Frontal lobe FTLN patients were relatively likely to generate the relational response for those picture pairs in which perceptual cues, such as spatial position, were correlated with the relational response, suggesting that their

overall relational deficit was due at least in part to impairment in the ability to inhibit the competing feature-based match. In contrast, temporal lobe FTLN patients were more uniformly impaired in relational responding across all picture pairs, suggesting that their underlying deficit involved loss of the conceptual knowledge required to encode the depicted relations.

The dissociation between the deficits of the two patient groups was revealed more clearly in Experiment 2, in which participants were asked to solve two-alternative verbal four-term analogy problems in which the association strength of the analogical response was greater than (positive SFI), equal to (neutral SFI), or less than (negative SFI) that of the foil. These problems can be solved on the basis of a single relation, and hence, even frontal lobe FTLN patients are capable of performing the necessary mapping and making the analogical response for positive and neutral SFI problems. However, frontal lobe FTLN patients were selectively impaired in solving the negative SFI problems, which require rapid learning of the analogical response so that its activation can be maintained, coupled with inhibition of the more highly associated foil. Temporal lobe FTLN patients, in contrast, were more uniformly impaired across all problem types, as would be expected if the source of their reasoning deficit involved neither analogical mapping nor inhibition, but rather loss of the underlying conceptual knowledge in long-term memory required to encode the relation between the A and B terms of the analogy. This finding is consistent with many studies of patients with temporal lobe FTLN that have found poor semantic memory performance in patients with damage to the anterior temporal cortex (Mummery et al., 1999; Hodges, Patterson, Oxbury, & Funnell, 1996).

The postulated deficits in the frontal and temporal lobe FTLN patient groups on the verbal analogy task were modeled within an integrated neurocomputational model of analogical reasoning, LISA (Hummel & Holyoak, 1997, 2003). We were able to simulate the observed pattern of frontal lobe deficits by impairing the rate of rapid learning of analogical connections, coupled with reduction of inhibitory control. Both rapid learning (Cools et al., 2002; O'Doherty et al., 2001; Rolls, 2000; Asaad et al., 1998) and inhibition (Miller & Cohen, 2001; Shimamura, 2000) appear to be key functions of the prefrontal cortex. When both these functions (not just either one alone) were impaired in LISA, the model yielded the pattern of selective impairment on negative SFI problems shown by frontal lobe FTLN patients. When the extent of semantic death (loss of connections between semantic units representing a relational role and a predicate unit for that role) was increased in LISA, thereby modeling a loss of conceptual knowledge in the anterior temporal cortex, the simulation yielded the pattern of impairment found for temporal lobe FTLN patients: a

relatively uniform decrease in accuracy across all verbal analogy problems, regardless of SFI condition.

In summary, the LISA simulations demonstrate that “computational lesioning” of rapid learning mechanisms and the inhibitory control of working memory yield deficits in analogical reasoning qualitatively similar to those observed with frontal lobe FTLN patients, whereas random destruction of connections within semantic memory yield deficits qualitatively similar to those observed with temporal lobe FTLN patients. The present study constitutes an initial step toward the goal of understanding the mechanisms of high-level human reasoning at a level that makes contact with what is known about cortical functioning.

METHODS

Experiment 1: Picture Analogy Task

Participants

Characteristics of participants in Experiment 1 are summarized in Table 1. Patients included in the study were diagnosed with FTLN according to the diagnostic criteria described in Neary et al. (1998). In addition, patients included in the study had abnormal scan data (SPECT, PET, or structural MRI; see Figure 1) and Mini-Mental Status Examination (MMSE) scores over 15. There were no reliable differences between patient groups and matched controls with respect to age and educational attainment. Patients and controls were paid US\$10 in exchange for participation in a one-hour testing session.

Materials

Picture analogy problems consisted of 10 pairs of pictures, eight provided by Markman and Gentner (1993) and two by Tohill and Holyoak (2000). Each of the pictures showed a visual scene involving three or more objects with at least one relation linking two of the objects (see Figure 2). In each problem, the same relation was present in both pictures; however, at least one of the objects linked by the relation was different in the two pictures (e.g., *restrain* [man, dog]; *restrain*

[tree, dog]). One object (e.g., man) in the top picture could be matched to an object in the bottom picture based either on a shared relational role (e.g., tree) or featural overlap (e.g., boy). Each problem appeared on a single sheet of paper.

Procedure

Participants were tested with a modification of Markman and Gentner’s (1993) “one-map” procedure. Participants studied each problem for 10 sec. After this study period, the experimenter pointed to the critical object in the top picture and asked the participant to point to one object in the bottom picture that “goes with” the indicated object in the top picture. Participants indicated their choice by pointing or saying the name of the object. Problems were presented in a fixed order randomly determined before the study. There were no additional instructions given to the participant and there were no example problems.

Experiment 2: Verbal Analogy Task

Participants

Characteristics of participants in Experiment 2 are summarized in Table 2. Patients were selected using the same criteria as in Experiment 1 and also received US\$10 for participating in a testing session. Word association ratings were collected from 150 undergraduate students at the University of California, Los Angeles (UCLA) who received extra course credit in exchange for participating. Verbal analogy problems were evaluated for difficulty by 54 UCLA undergraduate students who also received course credit in exchange for participating.

Materials

The verbal analogies were selected from a set of 200 two-choice verbal analogy problems provided by Michael Gardner (Sternberg & Nigro, 1980). The problems were of the type A:B:C:D or D’ where participants must select between choices D and D’ based on the relationship between A and B. The correct relation pair is designated C:D and the distractor pair is C:D’. Each pair of words (A:B, C:D, and C:D’) were related by one of five common relations (same, opposite, function, linear ordering, category). Word association ratings for the C–D and C–D’ pairs were collected from 150 UCLA undergraduates. Participants were asked to rate “how associated” the individual word pairs were using a 5-point scale. These values were normalized and an SFI was calculated for each problem. The SFI was defined as the difference between the z-score of the word association for the correct pair (C:D) minus the incorrect pair (C:D’). Thus, when the SFI is positive for a problem, semantic associations favor the correct

Table 1. Characteristics of Patient and Control Participants for Experiment 1: Picture Analogy

Group	Number of Participants	Mean Age	Mean Education	Mean MMSE
Frontal lobe FTLN patients	7	64.7	15.9	25.7
Temporal lobe FTLN patients	6	68.3	17.8	24.3
Control participants	8	66.2	15.8	na

Table 2. Characteristics of Patient and Control Participants for Experiment 2: Verbal Analogy

Group	Number of Participants	Mean Age	Mean Education	Mean MMSE
Frontal lobe FTLD patients	7	64.7	15.9	25.7
Temporal lobe FTLD patients	5	67.6	17.5	25
Control participants	10	64.2	16.4	na

analogical response. When the SFI is negative, semantic associations between C and D' favor the incorrect analogical response. Based on their SFI values, problems were divided into positive (SFI > .3), negative (SFI < -.3) and neutral (-.3 ≤ SFI ≤ .3) SFI groups. These problems were normed for difficulty on 54 UCLA students, and 24 problems were chosen based on SFI, accuracy and correct RT.

Procedure

Problems were presented one at a time using a notebook computer. An instruction screen containing one example was read to each participant and they had the opportunity to ask questions. After completion of a practice session of five problems, the experimenter began the task. After the participant indicated their choice either verbally or by pointing, the experimenter recorded their choice on the computer. Problems were presented in random order to each participant.

Simulation of Verbal Analogy Results in LISA

LISA represents propositions using a hierarchy of units (see Figure 5 for a schematic representation of LISA's architecture). At the bottom of the hierarchy, "semantic" units (small circles in Figure 5) represent objects and relational roles in a distributed fashion. Consider the proposition *becomes* (prince, king). Each role of the *becomes* relation would be represented by units coding for its semantic content (e.g., *prior-state* for the first role, *end-state* for the second, and *transition* for both).

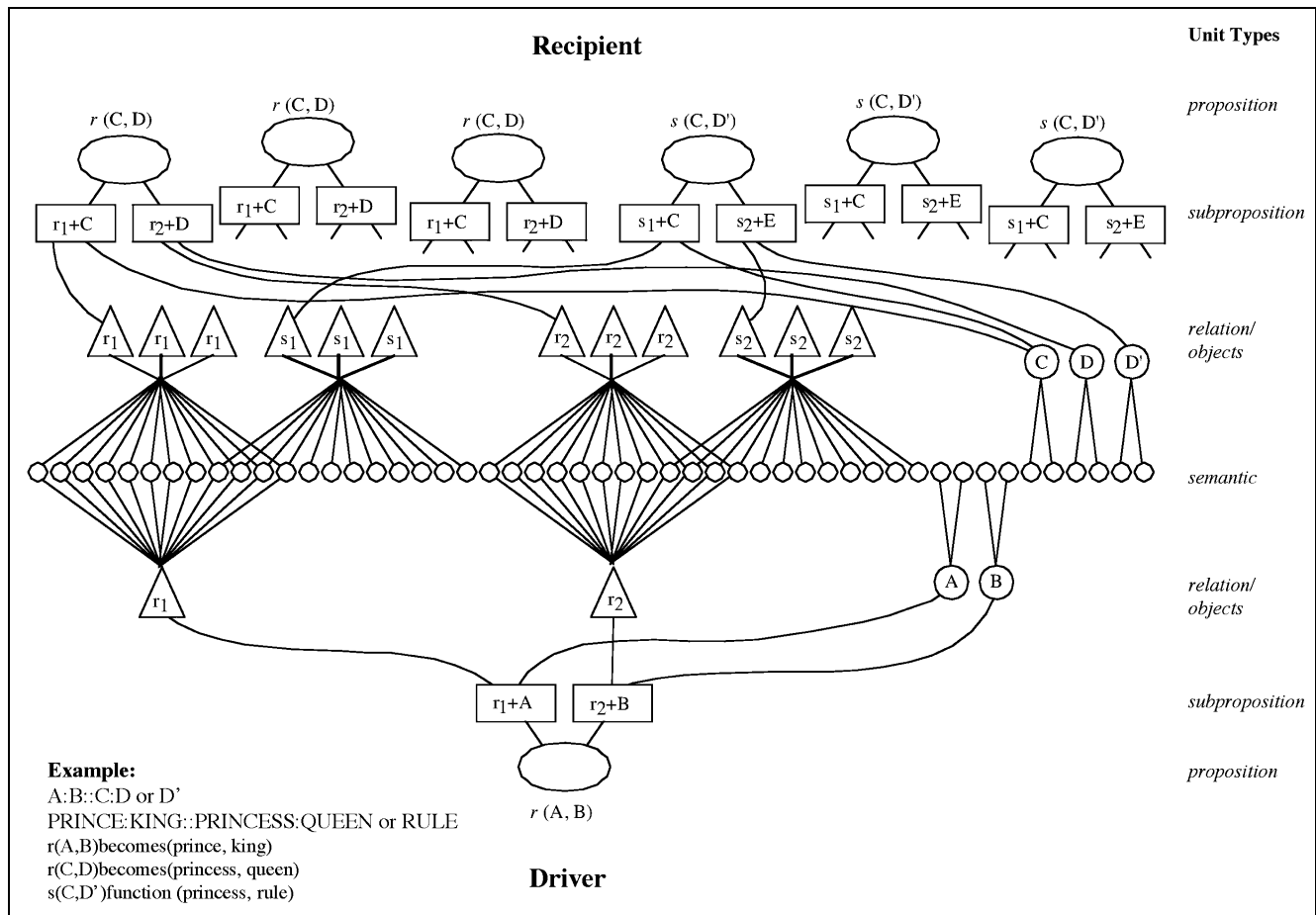


Figure 5. Modeling the verbal analogy task (the A:B:C:D or D' analogy problem) in LISA. The driver is represented by a set of hierarchical units that encode objects A and B that are related by relation r. The recipient is represented by sets of units that encode the correct relational response (D) as related to C by relation r. The competing response is also represented in the recipient, but by units associating C and D' by the relation s. The ratio of proposition units in the driver for the r(C:D) and s(C:D') is varied according to the SFI type of the problem.

Similarly, the arguments “prince” and “king” would be represented by units specifying their meaning (e.g., *person, male, royalty*). *Predicate* and *object* units (triangles and large circles, respectively, in Figure 5) represent relational roles and their fillers in a localist fashion, and have bidirectional excitatory connections to the corresponding semantic units. Subproposition (SP) units (rectangles in Figure 5) bind roles to their arguments, and have bidirectional connections to the corresponding predicate and object units. In the case of *becomes* (prince, king), one SP would bind “prince” to the first role of *becomes*, and another would bind “king” to the second. At the top of the hierarchy, proposition (P) units bind role-filler bindings into complete propositions via excitatory connections to the corresponding SPs. A complete analog (i.e., situation, story or event) is represented by the collection of semantic, predicate, object, SP and P units that collectively code the propositions in that analog. Separate analogs do not share object, predicate, SP or P units. However, all analogs are connected to the same set of semantic units. The semantic units thus permit the units in one analog to communicate with the units in others.

For the purposes of memory retrieval and analogical mapping (Hummel & Holyoak, 1997) as well as analogical inference and schema induction (Hummel & Holyoak, 2003), analogs are divided into two mutually exclusive sets: a *driver* and one or more *recipients*. The sequence of events is controlled by the driver: One (or at most three) at a time, propositions in the driver become active (i.e., enter working memory). When a proposition enters working memory, the binding of its roles to their arguments is represented by synchrony of firing: All the units under a given SP fire in synchrony with one another, and separate SPs fire *out* of synchrony with one another. The result on the semantic units is a set of mutually desynchronized patterns of activation: One pattern for each active SP (i.e., role binding) in the driver. In the case of *becomes* (prince, king), the semantic features of “prince” would fire in synchrony with the features of the first role of *becomes*, while “king” fires in synchrony with the second. To represent the proposition *becomes* (king, prince), LISA would activate exactly the same semantic units, but their synchrony relations would be reversed, with “king” firing in synchrony with the first role of *becomes*, and “prince” firing with the second. The resulting patterns of activation on the semantic units drive the activation of propositions in the various recipient analogs, and serve as the basis for analogical mapping, inference, schema induction, and the other functions LISA performs (Hummel & Holyoak, 1997, 2003). The final component of the LISA architecture is a set of *mapping connections* between units of the same type (e.g., object, predicate, etc.) in separate analogs. These connections grow whenever the corresponding units are active simultaneously, and

thereby permit LISA to learn the correspondences between structures in separate analogs.

To model the verbal analogies solved by the various patient groups in this study, we made several assumptions about the way people represent and solve these problems. In the positive SFI condition, there are more instances of relations matching A:B and C:D (the correct response) in long-term memory than there are instances of relations matching C:D' (the incorrect choice); in the negative SFI condition, there are more instances matching the incorrect choice; and in the neutral SFI condition, the two choices are matched. In the simulations, the numbers of relations assumed was 4:2, 2:4 and 3:3, respectively. The task involves mapping the A:B pair into long-term memory to find a match. That is, *Relation 1* (A, B) serves as the driver, and the various *Relation 1* (C, D) and *Relation 2* (C, D') propositions are in the recipient. Finding a match, that is, deciding whether the correct answer is D or D', is based on monitoring the units representing D and D' to determine which becomes more active.

Damage to the anterior temporal cortex corresponds in LISA to a loss of connections between the semantic properties of relational roles (i.e., semantic units) and units representing those roles in long-term memory. It was assumed that all types of participants (including normal controls) suffer from some amount of “noise” in their semantic representation of the situation at hand. This was modeled as a random input (in the range 0...*r*) added to the inputs to the semantic units at each instant in time. For reference, a semantic unit receives an input of approximately 1.0 from any driver or recipient unit to which it is connected, so the value of *r* represents a proportion of this (nonrandom) input. For these simulations the value of *r* was .33 for all conditions.

We also assumed that some proportion, *p*, of the connections from semantic units to units representing tokens of relational roles (predicate units) and to object fillers (object units) are randomly lost (“semantic death”). To simulate normal controls and frontal lobe FTLD patients, we set *p* to 0.1 (i.e., LISA forgets 10% of the semantic connections in these conditions); to simulate temporal lobe FTLD patients, *p* was set to 0.75. That is, we assume temporal lobe FTLD patients have lost much of their semantic knowledge of what the various relations mean.

We assume one cost of frontal degeneration is a general loss of the ability to inhibit irrelevant information. Accordingly, to simulate the frontal lobe FTLD patients' performance, we reduced ability of units in the recipient analogs to inhibit one another to 0.4; to simulate controls and temporal lobe FTLD patients, this parameter was left at its default value of 1.0. We assume that another function of the frontal lobes is to maintain a representation of the correspondences discovered during analogical mapping (i.e., the rapid learning of new

associations). Accordingly, to simulate frontal lobe FTLD patients, we reduced the mapping connection learning rate from 0.9 (its default value in LISA) to 0.4. To simulate controls and temporal lobe FTLD patients, this parameter was left at its default value of 0.9.

Acknowledgments

This work was supported by grants IBN-998-5417 from the National Science Foundation (BJK) and NRSA Training Grant MH-64244-01A1 (RGM). We thank Art Markman for providing the picture analogy problems and Michael K. Gardner for providing the verbal analogy problems. We would especially like to thank our patients and their caregivers who graciously participated in our study at a difficult time in their lives.

Reprint requests should be sent to Robert G. Morrison, Department of Psychology, University of California, Los Angeles, 405 Hilgard, Los Angeles, CA 90095-1563, or via e-mail: morrison@psych.ucla.edu.

REFERENCES

- Asaad, W. F., Rainer, G., & Miller, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron*, *21*, 1399–1407.
- Baddeley, A., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of working memory. *Quarterly Journal of Experimental Psychology*, *51A*, 819–852.
- Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review*, *10*, 48–71.
- 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.
- Cools, R., Clark, C., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 4563–4567.
- D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: Evidence from event related fMRI studies. *Experimental Brain Research*, *133*, 3–11.
- Dagenbach, D. E., & Carr, T. H. (1994). *Inhibitory processes in attention, memory, and language*. San Diego, CA: Academic Press.
- Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence*, *15*, 157–173.
- Dempster, F. N. & Brainerd, C. J. (Eds.). (1995). *Interference and inhibition in cognition*. San Diego, CA: Academic Press.
- Diamond, A. (1990). The development and neural bases of memory functions as indexed by the A-not-B and delayed response tasks in human infants and infant monkeys. In A. Diamond (Ed.), *The development and neural bases of higher cognitive functions* (pp. 267–317). New York: New York Academy of Sciences.
- French, R. M. (2002). The computational modeling of analogy-making. *Trends in Cognitive Science*, *6*, 200–205.
- Gentner, D., Holyoak, K. J., & Kokinov, B. N. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge: MIT Press.
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, *10*, 277–300.
- 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. *Behavioral and Brain Sciences*, *21*, 803–864.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.
- Hodges, J. R. (2000). Memory in the dementias. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 441–459). New York: Oxford University Press.
- Hodges, J. R., & Miller, B. (2001). The classification, genetics and neuropathology of frontotemporal dementia. Introduction to the special topic papers: Part I. *Neurocase*, *7*, 31–35.
- 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., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition*, *15*, 332–340.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*, 295–355.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge: MIT Press.
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*, 427–466.
- Hummel, J. E., & Holyoak, K. J. (2003). A symbolic–connectionist theory of relational inference and generalization. *Psychological Review*, *110*, 220–264.
- Kane, M. J., & Engle, R. W. (2003). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, *9*, 637–671.
- Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of dorsolateral and anterior prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, *12*, 477–485.
- Luo, Q., Perry, C., Peng, D., Xu, D., Ding, G., & Xu, S. (in press). The neural substrate of analogical reasoning: An fMRI study. *Brain Research*.
- Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, *25*, 431–467.
- Markowitsch, H. J. (2000). Neuroanatomy of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 465–484). New York: Oxford University Press.
- Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649–652.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
- 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.
- Mummary, C. J., Patterson, K., Wise, R. J. S., Vandenberg, R., Price, C. J., & Hodges, J. R. (1999). Disrupted temporal lobe connections in semantic dementia. *Brain, 122*, 61–73.
- Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P., Albert, M., Boone, K., Miller, B., Cummings, J., & Benson, D. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology, 51*, 1546–1554.
- O'Doherty, J., Kringelbach, M.L., Rolls E.T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*, 95–102.
- Prabhakaran, V., Smith, J. A. L., Desmond, J. E., & Glover, G. H. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology, 33*, 43–63.
- Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence, individual form*. London: Lewis.
- 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: MIT Press.
- Rolls, E. T. (1996). The orbitofrontal cortex. *Philosophical Transactions of the Royal Society London, B, 351*, 1433–1444.
- Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10*, 284–294.
- Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 629–639.
- Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 456–468.
- Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology, 28*, 207–218.
- Snow, R. E., Kyllonen, C. P., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 47–103). Hillsdale, NJ: Erlbaum.
- Spearman, C. (1923). *The nature of intelligence and the principles of cognition*. London: Macmillan.
- 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 Development, 51*, 27–38.
- Tohill, J. M., & Holyoak, K. J. (2000). The impact of anxiety on analogical reasoning. *Thinking & Reasoning, 6*, 27–40.
- Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., Thomas, C. R., & Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10*, 119–125.
- Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory and Cognition, 28*, 1205–1212.
- Welsh, M. C., Satterlee-Cartmell, T., & Stine, M. (1999). Towers of hanoi and london: Contribution of working memory and inhibition to performance. *Brain and Cognition, 41*, 231–242.
- Wharton, C. M., Grafman, J., Flitman, S. S., Hansen, E. K., Brauner, J., Marks, A., & Honda, M. (2000). Toward neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology, 40*, 173–197.