Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance

Soohyun Cho, Keith J. Holyoak, and Tyrone D. Cannon
University of California, Los Angeles, California

We report a series of experiments using a pictorial analogy task designed to manipulate relational integration, interference resolution, and active maintenance simultaneously. The difficulty of the problems was varied in terms of the number of relations to be integrated, the need for interference resolution, and the duration of maintenance required to correctly solve the analogy. The participants showed decreases in performance when integrating multiple relations, as compared with a single relation, and when interference resolution was required in solving the analogy. When the participants were required to integrate multiple relations while simultaneously engaged in interference resolution, performance was worse, as compared with problems that incorporated either of these features alone. Maintenance of information across delays in the range of 1–4.5 sec led to greater decrements in visual memory, as compared with analogical reasoning. Misleading information caused interference when it had been necessarily attended to and maintained in working memory and, hence, had to be actively suppressed. However, sources of conflict within information that had not been attended to or encoded into working memory did not interfere with the ongoing controlled information processing required for relational integration. The findings provide evidence that relational integration and interference resolution depend on shared cognitive resources in working memory during analogical reasoning.

A range of empirical evidence has established that analogical reasoning depends on the *executive* resources of working memory (WM). For example, experiments in which a dual-task methodology has been utilized have shown that the processes of binding and mapping used in analogical reasoning suffer interference from a concurrent WM task, especially a task that demands executive control (Morrison, Holyoak, & Truong, 2001). The prefrontal cortex, which plays a critical role in the executive aspects of WM (Glahn et al., 2002), has also been shown to mediate analogical reasoning. Wharton et al. (2000) used positron emission tomography to investigate the brain activation associated with solving geometric analogy problems. Regions active during analogical reasoning, but not during literal visual comparisons, included the left dorsomedial prefrontal cortex. Other studies have used paradigms similar to the Raven's Progressive Matrices Test (RPM; Raven, 1938), a task that requires relational reasoning. Waltz et al. (1999) found that frontal lobe patients (as compared with normal controls and patients with anterior temporal lobe damage) showed a marked deficit in RPM-type problems that required integration of multiple relations (Halford, 1993; Halford, Wilson, & Phillips, 1998; Robin & Holyoak, 1995). Several studies in which functional magnetic resonance imaging has been used have shown bilateral prefrontal cortex activation during RPM problem solving, with progressively greater involvement of the right dorsolateral and anterior prefrontal cortex at higher levels of relational complexity (Christoff et al., 2001; Kroger et al., 2002; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997).

Inhibition is also considered to be an important function of the executive component of working memory. Many researchers have viewed inhibition as an important function of the prefrontal cortex (for reviews, see Kane & Engle, 2003; Miller & Cohen, 2001; Shimamura, 2000). At the behavioral level, inhibition can refer to a broad range of cognitive processes, including withdrawing attention, abstaining from performing mental operations, and overriding a prepotent but goal-irrelevant response tendency. The nature and mechanism of these different types of inhibition may differ among one another. Solving analogies is likely to utilize inhibitory mechanisms to restrain processing of goal-irrelevant information in WM and to resolve conflict in problem solving. Behavioral studies of analogical mapping have shown that the ability to identify relational correspondences is impaired when competing featural matches are present (e.g., Gentner & Toupin, 1986; Holyoak & Koh, 1987; Richland, Morrison, & Holyoak, 2006; Ross, 1987, 1989). Patients with frontal damage show selective impairment on analogy problems that require suppression of feature-based alternative responses (Morrison et al., 2004).

S. Cho, soohyun@psych.ucla.edu

Although empirical research has implicated both relational integration and need for inhibition as WM demands on analogical reasoning, very few researchers have attempted to manipulate both of these factors at once, and to our knowledge, in no study has this been done while the possible additional role of short-term maintenance was examined. Moreover, previous studies have not considered how different types of inhibitory process engaged at different phases of information processing during analogical reasoning may differ in their natures and mechanisms.

Overview of the Present Study

Analogical reasoning provides an excellent vehicle for assessing the possible interactions among multiple demands on executive control, because a variety of factors can be varied within a single analogy problem. In the present study, using an analogical reasoning task varying relational complexity, need for conflict resolution, and active maintenance, we examined the degree to which these component processes of problem solving share a common capacity-limited pool of executive resources. Since both relational integration and interference resolution are processes that require the executive functions of WM and cognitive control, we hypothesized that there would be an overlap in the cognitive mechanisms underlying relational integration and interference resolution. We did not have an explicit hypothesis as to whether maintenance in WM would also share the same cognitive resources as those required by relational integration and interference resolution. We explored these issues by examining the effect of increasing the demand for each component process on behavioral performance. A simultaneous increase in demand for processes that depend on the same mechanism

is expected to result in multiplicative, rather than additive, effects on performance (S. Sternberg, 1969).

We examined the performance of college students on the people pieces analogy (PPA) task (R. J. Sternberg, 1977). This task was adapted by Morrison et al. (2001) and Viskontas, Morrison, Holyoak, Hummel, and Knowlton (2004) to systematically vary relational complexity orthogonally with the need for interference resolution while maintaining a constant level of visual complexity (see Figure 1). The PPA task, like other analogy tasks, requires mapping the relational structure in one situation onto another. Each term of the four-term analogy problem (A:B :: C:D) in the PPA task consists of a cartoon character that has one value on each of four binary traits (clothing color, gender, height, and width). On each trial, a subset of traits (one to four) was randomly chosen as specifying goal-relevant dimensions to which participants were asked to attend. Using only the goal-relevant traits as dimensions for comparison, the participants were asked to compare whether the two people in each pair had the same value of a given trait and then to determine whether the two within-pair relationships matched across pairs. If any of the attended within-pair relations did not match across pairs, the participants were to respond *different*, whereas if all of the attended within-pair relations matched, they were to respond *same*.

In almost all previous studies of reasoning with four-term analogy problems, temporal maintenance requirements have been minimized by presenting all four terms together and keeping the entire problem visible until the participant responded (for an exception, see R. J. Sternberg, 1977). Using this procedure with the PPA task, Viskontas et al. (2004) found that college students re-
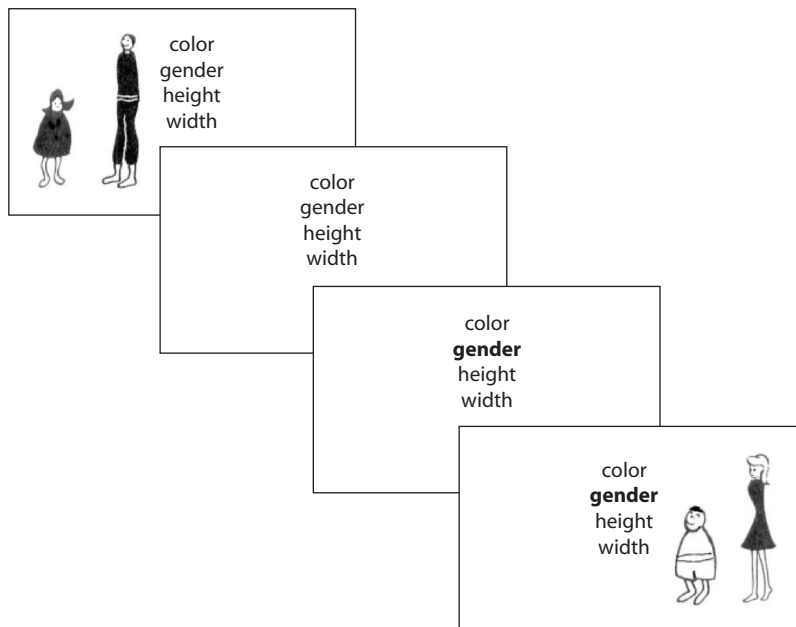

Figure 1. Example of analogy problem and task sequence used in Experiment 1. Bold font indicates relevant traits (colored red in the actual experiment; font size increased for legibility).

sponded more slowly when relational complexity was high but that response time (RT) was not affected by the number of dimensions that demanded conflict resolution. In contrast, the results for middle-aged and older adults showed an interaction between the two variables, so that increased need for conflict resolution slowed responding when relational complexity was high.

The results in Viskontas et al. (2004) seem to imply that young adults can solve complex analogy problems as accurately and as quickly, even when simultaneously engaged in conflict resolution. This finding challenges the hypothesis that relational integration and interference resolution jointly tax a common pool of executive resources in the intact human brain. However, two aspects of the experimental methods used by Viskontas et al. allow for alternative explanations for the pattern of findings observed in young adults. First, the list of relevant traits was shown from the very beginning of each trial, which enabled the participant to ignore goal-irrelevant information at encoding, thereby preventing possible sources of conflict from being encoded into WM and, therefore, minimizing the need to engage in conflict resolution (see Nee & Jonides, 2004, 2007). Second, as in almost all previous studies of reasoning with four-term analogy problems, active maintenance requirements were minimized by presenting all four terms of the analogy together and keeping the entire problem visible until the participant responded. Since the PPA task uses binary dimensions for which the relationship between A:B or C:D is either *same* or *different*, the participants could solve the analogy by verbally rehearsing the relations (the word *same* or *different*) and matching these verbal labels across pairs. Therefore, since all the terms were present on the screen, the participants may have relied more on the phonological loop than on the central executive of WM.

In order to make the PPA task more dependent on executive processes of *relational integration* and *interference resolution*, two features of the PPA task were modified in the present study. First, we modified the method in which the participants were cued to the goal-relevant information so that sources of conflict within goal-irrelevant information would not be filtered out at encoding. This was done by delaying the cue to the goal-relevant dimensions until the end of the period during which the participants would have encoded all stimuli into WM while not knowing the goal-relevance of stimulus dimensions. Second, in order to implement this *delayed-cuing* paradigm, a delay period was inserted between the A:B and the C:D pairs. The modified PPA task with delayed cuing was implemented as follows. Only the first two terms (A:B) of the analogy problem were presented initially; this pair then disappeared. After a variable delay, the goal-relevant dimensions were cued, after which the second pair (C:D) appeared, at which point the analogy could be solved. Note that since the participants did not know in advance which traits were goal relevant, they had to attend to and maintain all information about A:B in WM until they were cued as to the number and identity of the goal-relevant dimensions. Since it is difficult to verbally rehearse all possibly relevant information about A:B (based on four dimensions), this procedure renders verbal rehearsal an in-

effective strategy to use. When cued with the goal-relevant dimensions after the delay, the participants had to suppress processing of information that had been attended to and maintained in WM a moment ago. On half of the trials, the suppressed information was designed to cause interference, whereas on the other half, it was not. If interference resolution taxes the same limited-capacity resources in WM as does relational integration, even in healthy young adults, this design should allow a more sensitive basis on which to detect such interactive influences between these two executive processes. Note that this procedure requires participants to attend to and maintain the same amount of information in WM regardless of the relational complexity or the interference level of the trial. The distinction between *ignored* and *suppressed* information with regard to its possible influence on controlled information processing will be examined further in Experiment 3.

A secondary goal of the present study was to determine whether maintenance load differentially affects analogical reasoning and visual memory. In Experiment 2, analogy trials were intermixed with delayed matching-to-sample (DMS) trials that had the same structure and stimuli as the analogy problems. Wharton et al. (2000) had previously examined the effect of temporal delay on relational mapping based on analogical versus literal similarity. However, the simple geometrical shapes used in their study may have rendered the delay manipulation ineffective by allowing unwanted verbal and semantic rehearsal strategies for maintenance. Perhaps due to this limitation, the authors were not able to identify systematic effects of delay in either behavioral or neuroimaging data.

Accordingly, Experiment 2 was designed so that the participants would not benefit from verbal or semantic strategies during maintenance, which should enhance the effectiveness of the delay manipulation. On DMS trials, the participants had to judge whether the first pair of cartoon characters (A:B) was visually identical to the second pair presented. The participants were cued as to whether the trial was an analogy or a memory problem at the end of the delay interval, shortly before the second pair appeared. These interspersed memory trials were intended to encourage the participants to encode and maintain a visual representation of the first pair over the delay interval. At the same time, it is possible that a relational representation of the A:B pair could also be extracted from the visual inputs. A relational representation, which could be more directly useful in solving an analogy problem, may be less susceptible to temporal decay, as compared with a visual representation. Accordingly, we predicted that an increase in delay interval would be less detrimental to analogical judgments than to visual memory for the A:B pairs.

EXPERIMENT 1

Method

Participants. We recruited 67 undergraduate participants (16 of them male, 51 female) through the psychology department participant pool at the University of California, Los Angeles (UCLA). The participants were in the age range of 18–28 years. All except 2 participants were right-handed. The participants received course credit in return for their participation.

Design. Stimuli from the PPA task adapted by Morrison et al. (2001) were used in this experiment. In the present variant of the PPA task, each problem consisted of two pairs of human cartoon characters that could be described by four binary dimensions: clothing color (black or white), gender (male or female), height (tall or short), and width (wide or narrow) (see Figure 1). The participants' task was to determine whether or not the analogy was valid, on the basis of a subset of goal-relevant dimensions randomly selected for each trial. The participants were instructed to solve each problem on the basis of goal-relevant traits only and to inhibit solving the analogy on the basis of goal-irrelevant traits. The analogy was to be determined *true* if the relationship between A:B and C:D was the same regarding all relevant traits. If any of the across-pair relationships did not match for any of the relevant traits, the analogy was *false*.

The number of goal-relevant traits defined the level of relational complexity (RC level). The number of goal-irrelevant traits that had a potential to cause interference (by supporting an incorrect response) defined the level of interference (interference level). The level of interference was manipulated independently of RC level. The total number of traits was always four (the sum of the number of goal-relevant and goal-irrelevant traits). The number of relevant traits could vary from one to four, whereas the number of goal-irrelevant traits varied from zero to three, depending on the number of goal-relevant traits. Interference level was manipulated only for the *true* analogy trials. The level of interference could vary from Level 0 to 3 depending on RC level. (Note that at RC Level 4, a level of interference cannot be defined, since all four traits are goal relevant to the analogy.) For example, at RC Level 1, one trait is goal relevant, whereas the other three traits are goal irrelevant. At RC Level 2, two traits are goal relevant, whereas the other two are goal irrelevant. At Interference Level 0, the analogy was true not only when the relevant traits were taken into consideration, but also when the irrelevant traits were taken into consideration. Therefore, even if the participants mistakenly considered any irrelevant traits in solving the analogy, they would not be faced with conflict and, potentially, be misled to the wrong answer. However, at Interference Level 1, one of the goal-irrelevant traits had the potential to cause conflict and to mislead the participant to an incorrect response. In the example given in Figure 1, the relevant trait is *gender* (as indicated by bold font for "gender" in the trait list in the third time frame). The relationships for A:B and for C:D matched one another (*different* gender relationship for A:B equals *different* gender relationship for C:D); therefore, the correct response was *true*. However, a participant who had mistakenly used *color* in solving this problem would be led to conclude that the analogy was false, since the color relationships for pairs A:B and C:D did not match across pairs.

In Experiment 1, two levels of relational complexity (RC Levels 3 vs. 1) were contrasted. Need for interference resolution was defined at Levels 0 and 1. The length of delay determined the maintenance load (delay: 1,000 or 4,000 msec). The study design was thus a 2 × 2 × 2 repeated measures design with 12 true analogy (i.e., match) trials per condition for each participant (a total of 96 true trials). An additional 48 false analogy (i.e., nonmatch) trials were presented, but these were not included in the analysis for reasons that will be explained later in the Results section. The trials were completely counterbalanced regarding which subset of traits was selected to be goal relevant. We predicted that the participants would make more errors and take longer to solve problems involving higher levels of RC or interference resolution. Moreover, we hypothesized that a greater need for interference resolution (Interference Level 1) would especially interfere with analogical reasoning when the participants had to integrate three relations (RC Level 3) as compared with one (RC Level 1).

Procedure. Each trial started with a fixation cross located at the position at which the first pair would appear. After 500 msec, the first pair of cartoon people was presented along with the trait list. All trait names in the trait list were initially presented in black. The first pair disappeared after 1,000 msec, after which there was a variable temporal delay (700 or 3,700 msec), during which only the list

of trait names was presented. The participants were instructed to do their best to maintain a visual memory of the first pair without resorting to verbal rehearsal during this delay. The relevant traits in the trait list then turned red; 300 msec later, the second pair appeared at the right of the trait list. The participant was asked to decide whether the within-pair relations (same or different) for each relevant trait matched across pairs. The participant was told to press "1" if all of the within-pair relations matched across pairs and to press "2" if any of the relations did not match across pairs. The trial was terminated by the participant's self-paced keyboard press, and a new trial began after 1 sec.

The experiment was conducted on a desktop computer, and all the stimuli were presented on a CRT monitor. Stimuli generation and data collection were programmed with E-Prime software. The experiment lasted approximately 40 min. The participants were given short breaks during the experimental session.

Results and Discussion

RT and proportion correct (accuracy) were analyzed as dependent variables. For RT, geometric means were calculated for each participant in order to normalize variances. Only the data from correct trials were included in the analysis of RT. Data from participants whose accuracy was at or below chance (.5) in at least two experimental conditions (including at least one at RC Level 1) were excluded from the analysis of RT (6 of 67 participants). Also, correct trials with RTs outside the range of 650–15,000 msec were discarded from the analysis. There were 77 such outlier trials, which represented just 1.5% of the total number of correct trials. Only the trials in which all relevant across-pair relations matched (match trials) were included in the analysis, since trials with a nonmatching relationship were likely to differ in difficulty, depending on when the nonexhaustive search for a nonmatch was terminated.

The results, depicted in Figure 2, revealed that processing time increased with the need to integrate more complex relations (means of 2,266 msec for RC Level 1, 5,181 msec for RC Level 3) and to engage in interference resolution (3,638 msec for Interference Level 0, 3,808 msec for Interference Level 1) but was not affected by a longer duration of maintenance (3,737 msec for the 1,000-msec delay, 3,709 msec for the 4,000-msec delay). An ANOVA on the RT data revealed main effects of both relational complexity level [$F(1,60) = 428.01$, $MS_e = 2.42$, $p < .001$, $\eta^2 = .88$] and interference level [$F(1,60) = 13.30$, $MS_e = 0.26$, $p < .001$, $\eta^2 = .18$]. Moreover, a simultaneous increase in the load of relational integration and interference resolution increased processing time to a greater extent than could be accounted for by the additive effect of these factors [$F(1,60) = 19.68$, $MS_e = 0.18$, $p < .001$, $\eta^2 = .25$]. However, the interaction between the load of relational integration and maintenance was not significant [$F(1,60) = 1.93$, $MS_e = 0.23$, $p < .17$, $\eta^2 = .03$]. The main effect of delay, the two-way interaction between levels of interference and duration of delay, and the three-way interaction also were not significant from the analyses of RT.

Performance accuracy (see Table 1) decreased for more complex analogy conditions [.91 for RC Level 1, .81 for RC Level 3; $F(1,66) = 62.24$, $MS_e = 0.02$, $p < .001$, $\eta^2 = .49$] but was not affected by the need for interfer-
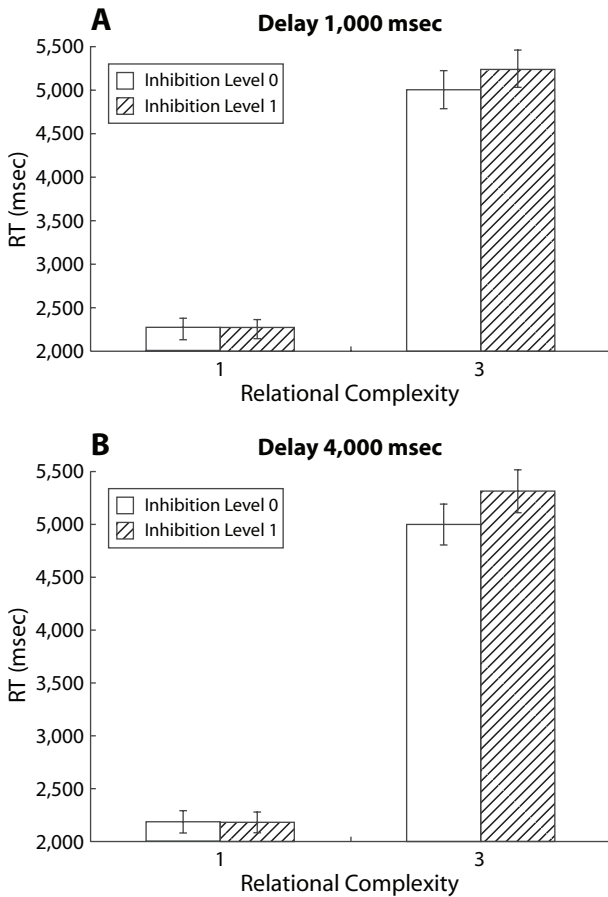
Figure 2. Mean response times (RTs) in Experiment 1 as a function of relational complexity (Level 1 vs. 3) and need for interference resolution (interference: Level 0 vs. 1). Bars indicate standard errors of the means.

ence resolution ($F < 1$) or by the duration of maintenance [$F(1,66) = 1.95$, $MS_e = 0.01$, $p = .17$, $\eta^2 = .03$]. Neither the two-way interactions nor the three-way interaction was significant in accuracy analyses.

In summary, the participants were slower when there was a need for interference resolution and integration of multiple relations. Processing speed deteriorated even further than would be predicted by additive effects of increased demand for interference resolution and relational integration when the load on both factors increased simultaneously. The participants also made more errors when integrating multiple relations, as compared with one.

The results of the present experiment suggest that even in young healthy individuals, integrating multiple relational information while resolving interference from conflicting information is a challenging task, performance on which deteriorates when the demands on relational integration and the need for interference resolution increase simultaneously.

Interestingly, the interdependency of relational integration and interference resolution was not affected by the length of the delay. One possible interpretation is that the representation required for analogical reasoning might

not be susceptible to decay, at least up to the temporal duration tested (4 sec). However, an alternative explanation for the lack of a delay effect might be that the design of Experiment 1 failed to completely prevent verbal rehearsal strategies during the delay. In order to block verbal rehearsal strategies that might minimize any impact of delayed cuing, Experiment 2 included visual memory trials, randomly intermixed among analogy trials, in order to force the participants to maintain visual representations of the A:B pair across the delay interval.

The present evidence of an interaction between levels of relational complexity and interference for young adults contrasts with the apparent lack of such an interaction in the study of Viskontas et al. (2004). The introduction of a delay prior to cuing the participants as to which traits were goal relevant may have been critical in generating the interaction, since it forced the participants to actively encode and maintain all the traits and then to selectively suppress goal-irrelevant information maintained in WM. Another possible explanation, however, is that the present experiment included a higher level of relation complexity (Level 3) than did that of Viskontas et al. (Level 2). Experiment 2 used RC Level 2 in order to determine whether an interaction between levels of RC and interference would still be found even when the maximal complexity level was the same as that in the earlier study.

EXPERIMENT 2

The goal of Experiment 2 was twofold. In order to determine whether the interaction between need for interference resolution and relational integration observed in Experiment 1 could be observed even at a lower complexity level, Experiment 2 used RC Levels 1 and 2 (instead of 1 and 3, as in Experiment 1). On the basis of evidence that RC Level 2 is the critical level at which integration of multiple relations is required (Waltz et al., 1999), we expected that we would again obtain an interaction with need for interference resolution.

Second, in order to examine the effect of temporal maintenance load on different types of representations, DMS (visual memory) trials were randomly intermixed with analogy trials in Experiment 2. The inclusion of visual memory trials was expected to prevent the participants from relying on verbal rehearsal of information about the first pair during the temporal delay. The structure of these memory trials was identical to that of the analogy

Table 1
Mean Proportions Correct and Standard Errors of the Means (*SEM*s) at Corresponding Levels of Each Independent Variable in Experiment 1

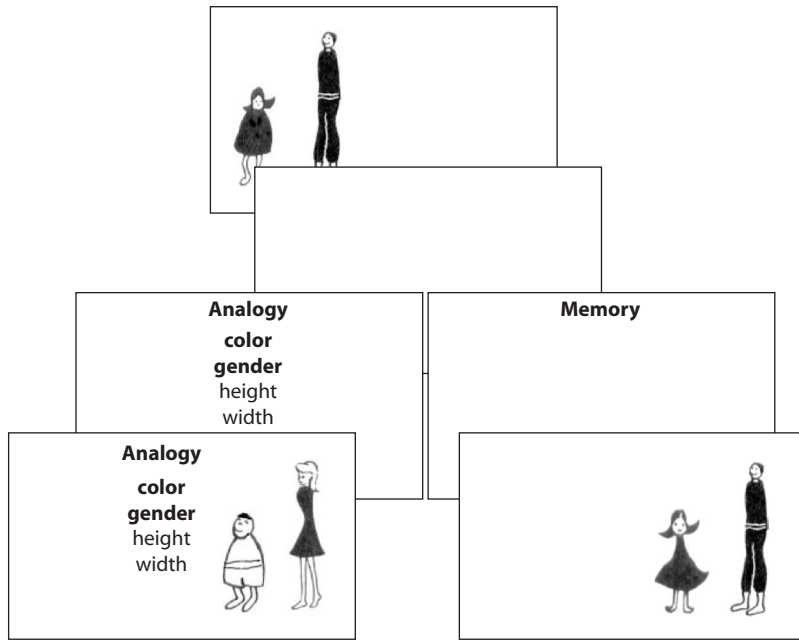| Delay (msec)	Level of Interference	Relational Complexity			
		Level 1		Level 3	
		M	*SEM*	*M*	*SEM*
1,000	0	.90	.02	.82	.02
	1	.92	.02	.81	.02
4,000	0	.90	.02	.80	.03
	1	.90	.02	.80	.02

Figure 3. Example of analogy trial task sequence (left) and delayed matching-to-sample memory trial task sequence (right) used in Experiment 2. Bold font indicates relevant traits (colored red in the actual experiment; font size increased for legibility).

trials. The only difference was that, after the delay, the participants were asked to determine whether or not the second pair (probe pair) was literally identical to the first pair (sample pair), which had been briefly presented prior to the delay. Since the participants did not know whether the current trial would be an analogy or a memory trial until the end of the delay, they would have to remember the visual stimulus as a whole (i.e., not just the extracted relational information) in order to be ready for both types of trials.

Method

Participants. We recruited 42 undergraduate participants (19 of them male and 23 female) through the psychology department at UCLA. The participants were in the age range of 18–25 years. All the participants except 1 were right-handed. The participants received course credit for 1 h in return for their participation.

Design and Procedure. The same set of stimuli as that in Experiment 1 was used in both the analogy and the memory trials in Experiment 2. Experiment 2 was designed to examine two hypotheses simultaneously: (1) whether relational complexity and interference level interact even when a contrast of RC Level 2 versus 1 is used and (2) whether the duration of maintenance differentially affects visual memory and analogical reasoning. Analogy trials were constructed by crossing relational complexity (RC level, 1 vs. 2) with levels of interference (interference level, 0 vs. 1) and delay between A:B and C:D pairs (delay: 1,500 vs. 4,500 msec). DMS memory trials were varied in terms of the duration of online maintenance (delay: 1,500 vs. 4,500 msec) inserted between the sample and the probe. When the effect of maintenance duration on analogy versus memory judgments was examined, only the analogy trials at RC Level 1 (collapsing over interference levels) were included in order to roughly match the difficulty of the two trial types. The set of analogy trials was completely counterbalanced with respect to which subset of traits was chosen to be goal relevant. The set of memory trials was also completely coun-

terbalanced as to which dimension was chosen to create a nonmatch. However, we did not use the entire counterbalanced list (which would have required over 75 min of problem-solving time), in order to prevent the participants from losing concentration. Accordingly, each participant solved 144 trials randomly chosen from the problem set, which resulted in an average of 8 match trials per condition.

We predicted that the participants would make more errors and would take longer to respond at the higher level of relational complexity (RC Level 2) and interference (Interference Level 1). Moreover, we expected that a need for interference resolution (Interference Level 1) would especially impair analogical reasoning when the participants had to integrate two relations (RC Level 2), as compared with one (RC Level 1). In addition, we predicted that increasing delay would be more detrimental to DMS judgments than to analogy judgments, since memory trials presumably rely more on a veridical visual representation of the A:B pair, whereas analogy trials may be based on a more abstract representation of relations between trait values.

Each trial (see Figure 3) started with a fixation cross located at the position where the first pair would appear. After 500 msec, the first pair of cartoon people was presented (without the list of trait names). The first pair disappeared after 1,500 msec, after which there was a variable temporal delay (1,000 or 4,000 msec), during which a blank screen was shown. The participants were instructed to do their best to maintain a visual memory of the first pair without resorting to verbal rehearsal during this delay. After the delay, the word "Analogy" or else "Memory" appeared at the top of the screen as an indication of trial type. On an analogy trial, the trait list then appeared, with the relevant traits colored red. No trait list was shown on memory trials. After another 500 msec, a second pair of cartoon people appeared at the location to the right of where the trait list appeared on analogy trials. On analogy trials, the second pair was the C:D pair, to complete the analogy problem; on memory trials, it was either a repetition of the A:B pair (*match* trials) or a pair in which the value of one trait in one character was different (*nonmatch* trials). Note that the total duration of delay (1,500 vs. 4,500 msec) was 500 msec longer than in Experiment 1.
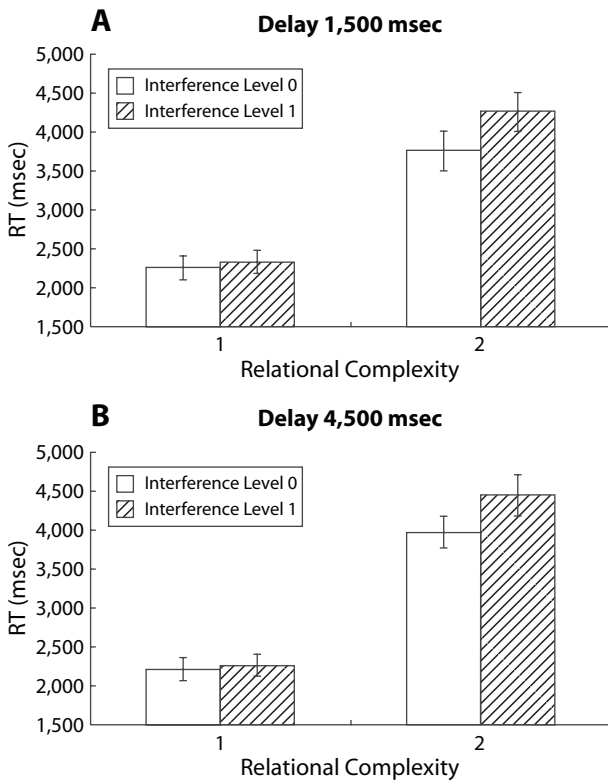
Figure 4. Mean response times (RTs) on analogy trials in Experiment 2 as a function of relational complexity (Level 1 vs. 2) and need for interference resolution (interference: Level 0 vs. 1). Bars indicate standard errors of the means.

On analogy trials, the participant was asked to decide whether the within-pair relations (same or different) for each goal-relevant trait matched across pairs. The participant was instructed to press "1" if there was any nonmatching relationship across pairs (on analogy trials) or if the probe pair did not match the sample pair (on memory trials) and to press "0" if all across-pair relationships matched (on analogy trials) or if the probe matched the sample (on memory trials). The participant's self-paced keyboard press terminated the trial. A new trial began after 1 sec. The experiment lasted approximately 35 min. The participants were given short breaks during the experimental session.

Results and Discussion

Data from the participants whose accuracy was at or below chance in at least two experimental conditions (including at least one condition at RC Level 1) on either analogy or memory trials were excluded from the analysis of RT (8 of 42 participants). RTs exceeding 15,000 msec on correct analogy or memory trials were also discarded from the analysis. There were eight such outlier trials, which constituted just 0.53% of the total number of correct trials. As in Experiment 1, RTs (based on geometric means) and proportions correct were analyzed as dependent variables. As in Experiment 1, only the trials on which the correct response was to report a match were included in the analyses.

Analogy judgments. The results, depicted in Figure 4, revealed that processing time increased with the need to integrate more complex relations (means of 2,271 msec

for RC Level 1, 4,115 msec for RC Level 2) and to engage in interference resolution (3,055 msec for Interference Level 0, 3,330 msec for Interference Level 1) but was not affected by a longer duration of maintenance (3,157 msec for the 1,500-msec delay, 3,228 msec for the 4,500-msec delay). An ANOVA on RT data revealed main effects of both relational complexity [$F(1,33) = 268.49$, $MS_e = 0.86$, $p < .001$, $\eta^2 = .89$] and interference level [$F(1,33) = 10.77$, $MS_e = 0.48$, $p < .002$, $\eta^2 = .25$]. Moreover, a simultaneous increase in the load of relational integration and interference resolution increased processing time to a greater extent than could be accounted for by the additive effect of these factors [$F(1,33) = 6.33$, $MS_e = 0.50$, $p < .02$, $\eta^2 = .16$]. The main effect of delay and all other interaction effects from RT analyses were not significant ($F < 1$).

Performance accuracy (proportion correct; see Table 2 for details) decreased for more complex analogy conditions [.86 for RC Level 1, .76 for RC Level 2; $F(1,41) = 30.98$, $MS_e = 0.02$, $p < .001$, $\eta^2 = .43$] but was not affected by the need for interference resolution [$F(1,41) = 2.48$, $MS_e = 0.02$, $p < .12$, $\eta^2 = .06$] or by the duration of maintenance [$F(1,41) = 1.52$, $MS_e = 0.01$, $p = .22$, $\eta^2 = .04$]. Simultaneous increases in levels of RC and interference resulted in significant decreases in accuracy [$F(1,41) = 5.54$, $MS_e = 0.02$, $p < .02$, $\eta^2 = .12$]. However, an overadditive effect of increasing RC and delay resulted in a significant decrease in accuracy [$F(1,41) = 6.36$, $MS_e = 0.01$, $p < .02$, $\eta^2 = .13$]. Neither the interaction effect between interference level and delay nor the three-way interaction was reliable ($F < 1$).

As was the case in Experiment 1, the participants in Experiment 2 made more errors and were slower when they were simultaneously engaged in interference resolution and integration of multiple relations. Considering the results of Experiments 1 and 2 together, it appears that the need to integrate multiple relations (RC Level 2 or above) interacts with the processing required to resolve interference from conflicting information.

Analogy versus memory judgments. As is depicted in Figure 5 and Table 3, performance on analogy trials at RC Level 1 (collapsing over interference levels) was compared with that on DMS memory trials at each delay interval. The participants took longer on analogy trials (2,269 msec) than on memory trials (1,790 msec) [$F(1,33) = 25.46$, $MS_e = 0.31$, $p < .001$, $\eta^2 = .44$] and were less accurate on analogy trials (.85) than on mem-

Table 2
Mean Proportions Correct and Standard Errors of the Means (*SEM*s) at Corresponding Levels of Each Independent Variable on Analogy Trials in Experiment 2

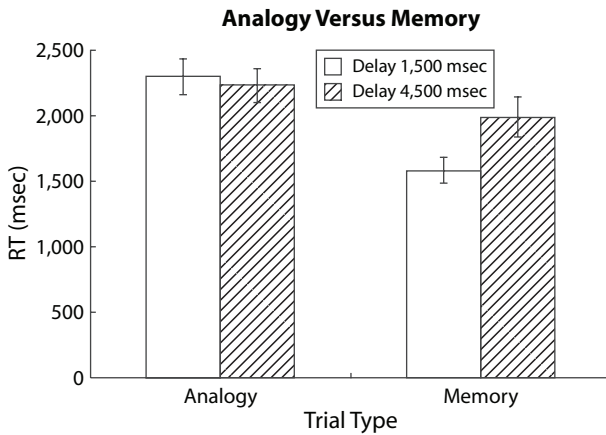| Delay (msec) | Level of Interference | Relational Complexity | | | |
| | | Level 1 | | Level 2 | |
		M	SEM	M	SEM
1,500	0	.88	.03	.78	.05
	1	.88	.03	.73	.04
4,500	0	.83	.03	.80	.04
	1	.84	.04	.73	.04

Analogy Versus Memory

Figure 5. Mean response times (RTs) in Experiment 2 as a function of trial type (analogy vs. memory) and delay (1,500 vs. 4,500 msec). Bars indicate standard errors of the means.

ory trials (.94) [$F(1,41) = 9.72$, $MS_e = 0.03$, $p < .003$, $\eta^2 = .19$]. Overall, an increase in the load of temporal maintenance caused performance to decline in terms of RT [$F(1,33) = 8.10$, $MS_e = 0.12$, $p < .008$, $\eta^2 = .20$] and accuracy [$F(1,41) = 16.47$, $MS_e = 0.005$, $p < .001$, $\eta^2 = .29$]. However, as is apparent from the RT pattern depicted in Figure 5, a longer delay increased processing time much more for memory than for analogy trials, resulting in a reliable interaction between delay and trial type [$F(1,33) = 19.03$, $MS_e = 0.10$, $p < .001$, $\eta^2 = .37$]. Performance accuracy of analogy and memory trials declined to a similar extent with an increase in duration of maintenance, so that the interaction between trial type and delay was not reliable in the analysis of accuracy.

The results of Experiment 2 indicate that imposing a longer delay impaired processing speed for memory judgments, but not for analogy judgments, even though memory decisions were made significantly more quickly than analogy decisions, regardless of delay. This pattern suggests that the duration of active maintenance impairs the quality of representations required for visual memory, but not so much for analogical reasoning. One possible interpretation is that the relational representations necessary for analogical reasoning are less susceptible to temporal decay than are the visual representations required for DMS recognition performance.

EXPERIMENT 3

A previous study by Viskontas et al. (2004) showed that relational complexity and interference resolution (in their terminology, "inhibition") created overadditive difficulties in analogical reasoning for older participants, but not for college students. Their findings would lead to a conclusion that for healthy young adults, relational integration and interference resolution do not compete for executive resources that are limited in capacity. However, the procedure used in their study made it possible for participants to solve the analogy by perceptually screening out goal-irrelevant dimensions of variation without atten-

tive, controlled processing. This type of inhibition, which operates by withdrawing attention from goal-irrelevant information, might require fewer executive resources, as compared with other types of inhibition that require suppressing information that is maintained in WM (Nee & Jonides, 2004, 2007); hence, no interaction effects were observed between relational integration and their manipulation of inhibition. In the present Experiments 1 and 2, we found that even healthy young adults encounter overadditive difficulty in integrating multiple relations in the presence of conflict caused by misleading information that had been attended to and maintained in WM. More specifically, two factors might have contributed to the greater impact of conflict-arousing information on analogical reasoning in the present study, in contrast to that in Viskontas et al.: (1) adding a requirement for maintenance of information about the A:B pair in WM and/or (2) delayed cuing of the goal-relevant dimensions until after the participants had actively processed information about the A:B pair in WM.

The purpose of Experiment 3 was to determine whether it is the delayed cuing of goal-relevant dimensions or the insertion of a temporal delay that is critical in generating an interaction between levels of relational complexity and of interference resolution. If delayed cuing is the critical factor, running the same procedure with early cuing of goal-relevant dimensions would not result in a significant main effect of interference level or any interaction effects involving the need for interference resolution. Following this logic, the participants in Experiment 3 were given cues to the goal-relevant traits from the beginning to the end of each trial, while the requirement for maintenance in WM was varied (need for maintenance: maintenance vs. no maintenance) randomly from trial to trial. With early cuing, the participants were not forced to attend to and maintain goal-irrelevant information in WM. Instead, the participants were able to screen out and ignore possibly misleading information without further processing from encoding and beyond.

Method

Participants. We recruited 10 participants (2 of them male and 8 female) through the psychology department at UCLA. The participants were in the age range of 18–29 years. All the participants were right-handed. All the participants, except for 1 volunteer, received course credit for 1 h in return for their participation.

Design and Procedure. The same set of stimuli as that in Experiments 1 and 2 was used in Experiment 3. The design was a $2 \times 2 \times 2$ repeated measures design that crossed need for maintenance (mainte-

Table 3
Mean Proportions Correct and Standard Errors of the Means
(*SEM*s) at Corresponding Levels of Each Independent
Variable in Experiment 2

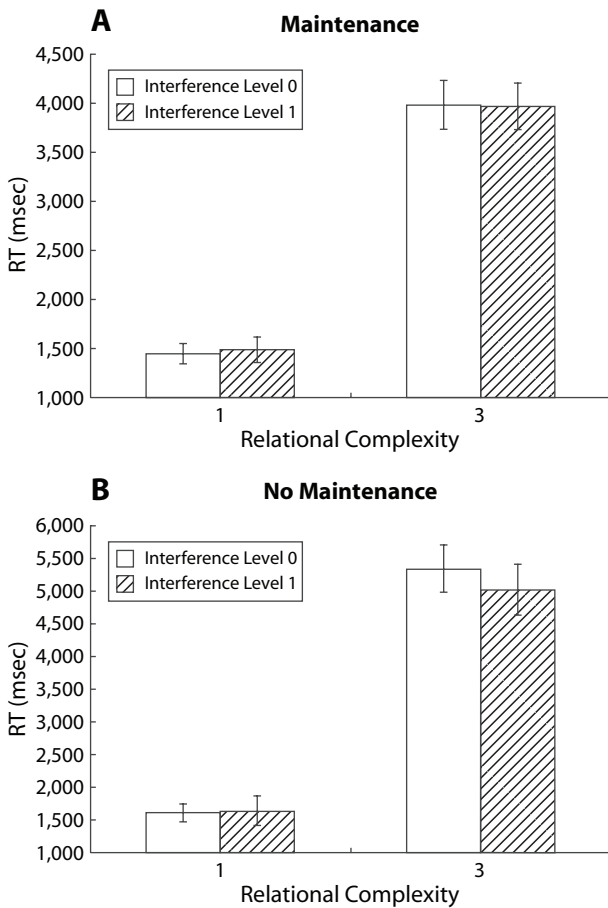| Delay (msec)	Trial Type			
	Analogy		Memory	
	M	SEM	M	SEM
1,500	.88	.03	.96	.01
4,500	.83	.03	.92	.02

A   Maintenance

B   No Maintenance

Figure 6. Mean response times (RTs) in Experiment 3 as a function of relational complexity (Level 3 vs. 1), need for maintenance (maintenance vs. no maintenance), and need for interference resolution (interference: Level 0 vs. 1). Bars indicate standard errors of the means. (A) Maintenance. (B) No maintenance.

nance vs. no maintenance), relational complexity (RC level: 1 vs. 3), and need for interference resolution (interference level: 0 vs. 1).

On 75% of the trials, the A:B pair disappeared after its initial presentation (maintenance trials), whereas on the remainder (25%) of the trials, the A:B pair stayed on the screen during the delay and throughout the trial (no-maintenance trials). Note that on no-maintenance trials, the participants would not be required to rely on WM to solve analogies. Three times as many maintenance trials as no-maintenance trials were included in order to encourage the participants to attend to and encode information about the first pair during the initial phase of each trial. Early cuing of goal-relevant traits would allow the participants to withdraw attention from traits that were cued to be goal irrelevant. Such a procedure was expected to reduce or eliminate any effects involving need for interference resolution that had been observed in Experiments 1 and 2. There were 18 match trials per condition for no-maintenance trials and 6 match trials per condition for maintenance trials.

At the beginning of each trial, the A:B pair was presented at the left side of the screen, along with the list of trait names presented at the center. The goal-relevant traits were presented in red, and goal-irrelevant traits were in black. The trials were counterbalanced with respect to which subset of traits was chosen to be goal relevant. The participants were given a maximum of 5 sec to encode and process the first pair. The participants were allowed to terminate the encod-

ing phase before the 5 sec were up by pressing the space bar. On maintenance trials, the A:B pair was removed for 1,000 msec, and then the C:D pair appeared on the screen. On no-maintenance trials, the A:B pair stayed on the screen during the delay for 1,000 msec, and then the C:D pair was added on the right side of the screen. (Note that no-maintenance trials did not have A:B and C:D pairs simultaneously presented during the encoding and delay phases.) Upon the appearance of the C:D pair, the participants made a response (self-paced), upon which the trial was terminated. The participants were asked to press "T" when analogical comparisons were true regarding every goal-relevant dimension, but otherwise, "N."

Results and Discussion

No participants were excluded from the analysis due to low performance accuracy. The results indicated that early cuing of goal-relevant traits, either with or without a requirement for visual WM, eliminated the effects of interference levels. There was no significant difference in either RT (3,099 msec for Level 0, 3,031 msec for Level 1) or accuracy (.95 for Level 0, .94 for Level 1) for different levels of interference. The effects of increasing relational complexity on RT [1,547 msec for RC Level 1, 4,583 msec for RC Level 3; $F(1,9) = 153.86$, $MS_e = 1.20$, $p < .001$, $\eta^2 = .95$] and on accuracy [.97 for RC Level 1, .92 for RC Level 3; $F(1,9) = 6.90$, $MS_e = 0.006$, $p < .03$, $\eta^2 = .43$] were similar to those observed in Experiments 1 and 2 (see Figure 6 for an analysis of performance RT and Table 4 for performance accuracy in each condition).

The results of Experiment 3 indicate that early cuing of goal-irrelevant traits, regardless of whether or not maintenance of the A:B pair is required, eliminates performance decrements due to an increase in interference level. (Note that behavioral performance in Experiment 3 was more accurate and faster than that in Experiment 1 at the same level of RC and interference, a finding that can be interpreted as indicating that early cuing of goal-relevant traits makes the PPA task easier for participants.)

These findings from Experiment 3, along with evidence from an additional follow-up experiment,[1] indicate that delaying the cue that informs the participant of the goal-relevant traits is a critical requirement for probing the process of interference resolution that competes with analogical reasoning for executive resources in healthy young adults. The present findings confirm that this methodological difference accounts for the robust interaction between levels of interference and relational complexity observed in Experiments 1 and 2, but not in the study of Viskontas et al. (2004).

Table 4
Mean Proportions Correct and Standard Errors of the Means (*SEM*s) at Corresponding Levels of Each Independent Variable in Experiment 3

Need for Maintenance	Interference Level	Relational Complexity			
		Level 1		Level 3	
		M	*SEM*	*M*	*SEM*
Maintenance	0	.95	.02	.91	.04
	1	.98	.01	.90	.03
No maintenance	0	.98	.02	.95	.04
	1	.95	.04	.92	.04

GENERAL DISCUSSION

The results of the present study provide the first behavioral evidence that both relational integration and interference resolution depend on a common pool of executive resources obtained by simultaneous manipulation of both factors within a single task in the young and healthy human population. Moreover, by differentiating between qualitatively different types of inhibition and probing the process of interference resolution that demands executive control, we found that even healthy young adults encounter overadditive difficulty in coping with integration of multiple relations while simultaneously engaged in resolution of interference. We also confirmed that interference, the resolution of which demands executive control, is caused only by information that has been attended to and is actively maintained in WM. In contrast, potential sources of interference based on information that has not been attended to and is filtered out during the encoding phase do not impair ongoing controlled information processing. A broader implication of our findings is that relational integration and inhibitory control over information that has been attended to and actively processed in WM likely depend on a shared neural substrate in the prefrontal cortex (Robin & Holyoak, 1995).

The other major finding from the present study, based on the results of Experiment 2, is that judgments of analogical similarity and of literal similarity in WM are differentially affected by the imposition of a maintenance requirement. In the present Experiment 2, a delay of varying length was always imposed between presentation of the A:B and the C:D items, and hence, information about A:B needed to be held in memory at the time of the analogical judgment. Imposing a longer delay caused decrements in the accuracy of both judgment types to a similar degree but caused processing time to increase only for memory judgments, even though such memory judgments were actually made more quickly overall than were the easiest type (RC Level 1) of analogy judgments. The fact that delay impaired the speed of memory judgments much more than it did the speed of analogy judgments may suggest that the two types of decisions were made on the basis of qualitatively distinct internal representations. The delayed recognition task likely depended on an image-like representation of the initial cartoon figures, and this quasiperceptual representation was impaired as the length of the delay was increased. In contrast, analogical judgments inherently depend on relational representations (e.g., the A and B cartoons were *same* in color but *different* in height). Maintenance of such relational representations may be more robust across delays of a few seconds. More generally, the human cognitive ability to manipulate information and solve complex problems may depend critically on the availability of symbolic representations that can be maintained in WM with minimal decay across prolonged intervals of processing.

AUTHOR NOTE

REFERENCES

CHRISTOFF, K., PRABHAKARAN, V., DORFMAN, J., ZHAO, Z., KROGER, J. K., HOLYOAK, K. J., & GABRIELI, J. D. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage*, **14**, 1136-1149.

GENTNER, D., & TOUPIN, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, **10**, 277-300.

GLAHN, D. C., KIM, J., COHEN, M. S., POUTANEN, V. P., THERMAN, S., BAVA, S., ET AL. (2002). Maintenance and manipulation in spatial working memory: Dissociations in the prefrontal cortex. *NeuroImage*, **17**, 201-213.

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 & Brain Sciences*, **21**, 803-865.

HOLYOAK, K. J., & KOH, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, **15**, 332-340.

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 & 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 anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, **12**, 477-485.

MILLER, E. K., & COHEN, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, **24**, 167-202.

MORRISON, R. G., HOLYOAK, K. J., & TRUONG, B. (2001). Working-memory modularity in analogical reasoning. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 663-668). Mahwah, NJ: Erlbaum.

MORRISON, R. G., KRAWCZYK, D. C., HOLYOAK, K. J., HUMMEL, J. E., CHOW, T. W., MILLER, B. L., & KNOWLTON, B. J. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience*, **16**, 260-271.

NEE, D. E., & JONIDES, J. (2004, November). *Dissociating information-suppression during encoding versus working memory*. Paper presented at the annual meeting of the Psychonomic Society, Minneapolis.

NEE, D. E., & JONIDES, J. (2007, May). *Dissociable neural mechanisms of interference-resolution: Perception and memory*. Poster presented at the annual meeting of the Cognitive Neuroscience Society, New York.

PRABHAKARAN, V., SMITH, J. A., DESMOND, J. E., GLOVER, G. H., & GABRIELI, J. D. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology*, **33**, 43-63.

RAVEN, J. C. (1938). *Progressive matrices: A perceptual test of intelligence, individual form*. London: Lewis.

RICHLAND, L. E., MORRISON, R. G., & HOLYOAK, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, **94**, 249-271.

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

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, & 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, & Cognition*, **15**, 456-468.

SHIMAMURA, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, **28**, 207-218.

STERNBERG, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.

STERNBERG, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, **30**, 276-315.

VISKONTAS, I. V., MORRISON, R. G., HOLYOAK, K. J., HUMMEL, J. E., & KNOWLTON, B. J. (2004). Relational integration, inhibition and analogical reasoning in older adults. *Psychology & Aging*, **19**, 581-591.

WALTZ, J. A., KNOWLTON, B. J., HOLYOAK, K. J., BOONE, K. B., MISHKIN, F. S., DE MENEZES SANTOS, M., ET AL. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, **10**, 119-125.

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.

NOTE

1. Additional support for the conclusion of Experiment 3 is provided by a further experiment using delayed cuing of goal-relevant traits while eliminating maintenance requirements. The structure of each trial was similar to that of the no-maintenance trials in Experiment 3. At the beginning of each trial, the A:B pair, along with the trait list (all trait names in black ink), were presented for 1,700 msec. Then the relevant dimensions in the trait list turned red while the A:B pair stayed on the screen. After 300 msec, the C:D pair was added onto the screen. The participants were asked to press "1" if the analogy was true for every relevant trait, but otherwise, "2." The participants were asked to make a response within 6 sec. If the participants did not answer within 6 sec, the trial was terminated. The results from 17 participants indicated that delayed cuing, even without a requirement for visual WM, induced significant increases in RT [2,824 msec for Level 0, 2,917 msec for Level 1; $F(1,16) = 4.95$, $MS_e = 0.03$, $p < .04$, $\eta^2 = .24$] and significant decreases in accuracy [.97 for Level 0, .88 for Level 1 $F(1,16) = 26.00$, $MS_e = 0.005$, $p < .001$, $\eta^2 = .62$], when demands for interference resolution increased. Also, as in Experiments 1 and 2, the increase in RT for a higher level of interference was accentuated at RC Level 3 versus 1, resulting in a significant interaction between RC and interference level [$F(1,16) = 11.85$, $MS_e = 0.02$, $p < .003$, $\eta^2 = .43$].