Navigating Increasing Levels of Relational Complexity: Perceptual, Analogical, and System Mappings

Matthew J. Kmiecik[1], Rodolfo Perez[1], and Daniel C. Krawczyk[1,2]

Abstract

■ Relational thinking involves comparing abstract relationships between mental representations that vary in complexity; however, this complexity is rarely made explicit during everyday comparisons. This study explored how people naturally navigate relational complexity and interference using a novel relational match-to-sample (RMTS) task with both minimal and relationally directed instruction to observe changes in performance across three levels of relational complexity: perceptual, analogy, and system mappings. Individual working memory and relational abilities were examined to understand RMTS performance and susceptibility to interfering relational structures. Trials were presented without practice across four blocks, and participants received feedback after each attempt to guide learning. Experiment 1 instructed participants to select the target that best matched the sample, whereas Experiment 2 additionally directed participants' attention to same and different relations. Participants in Experiment 2 demonstrated improved performance when solving analogical mappings, suggesting that directing attention to relational characteristics affected behavior. Higher performing participants—those with above-chance performance on the final block of system mappings—solved more analogical RMTS problems and had greater visuospatial working memory, abstraction, verbal analogy, and scene analogy scores compared to lower performers. Lower performers were less dynamic in their performance across blocks and demonstrated negative relationships between analogy and system mapping accuracy, suggesting increased interference between these relational structures. Participant performance on RMTS problems did not change monotonically with relational complexity, suggesting that increases in relational complexity places nonlinear demands on working memory. We argue that competing relational information causes additional interference, especially in individuals with lower executive function abilities. ■

INTRODUCTION

The foundation of human thinking and reasoning relies on relational thought. When we think relationally, we are able to understand and appreciate how mental representations relate to each other despite their featural discrepancies. For instance, consider an analogy between Ohm's law and water pressure. In an electrical circuit, Ohm's law ($V = IR$) states that voltage (V) is directly proportional to the amount of current (I) given a constant resistance (R), therefore implying that an increase in resistance, given a constant current, results in an increase in voltage. Similarly, the water pressure of a constant flow of water through a hose will increase if the hose becomes narrower. In this case, a relational mapping of concepts (water pressure-to-voltage; water flow-to-current; hose width-to-resistance) facilitates the understanding of their relationships, despite their featural differences (e.g., concepts like current, voltage, and resistance are impossible to see with the naked eye). Our ability to extract, compare, and integrate relationships enables myriad cognitive abilities ubiquitous in daily life (Hofstadter & Sander, 2013; Holyoak & Thagard, 1995; Gentner, 1983) that include learning (Vendetti, Matlen, Richland, & Bunge, 2015; Gentner, 2010; Richland, Zur, & Holyoak, 2007; Goswami, 1992), problem solving (Gick & Holyoak, 1980, 1983), creative thinking (Green, Kraemer, Fugelsang, Gray, & Dunbar, 2012), and discovery (Hofstadter & Sander, 2013; Gentner, 2002).

Analogical reasoning is a specific form of relational reasoning defined by the mapping of objects between knowledge structures that prioritizes relationships rather than attributes (Gentner, 1983). Object attributes and their degree of featural similarity are allowed to vary; however, the objects and their relations must maintain a systematic structure in which the relations between objects are the same. In essence, analogies involve the comparisons of relationships—water "flowing through" a hose is like current "flowing through" a circuit. Analogies can also be compared to one another. For instance, consider again the Ohm's law hydraulic analogy. There are often two different forms of this analogy offered that differ in how the water receives the energy necessary to flow through the hose either via (1) an electric water pump or (2) potential energy given to a bucket of water raised on an elevated surface. Despite these analogies not differing in their inherent relational structure, the reasoner in developing a

This article is part of a Special Focus, Relational Reasoning, deriving from a symposium at the 2019 annual meeting of the Cognitive Neuroscience Society, organized by Silvia Bunge and Keith Holyoak.

[1]The University of Texas at Dallas, [2]The University of Texas Southwestern Medical Center

Journal of Cognitive Neuroscience 33:3, pp. 357–376
https://doi.org/10.1162/jocn_a_01618

preference for either analogy, or even merely appreciating their relational similarities, has engaged in comparing higher-order relations (e.g., relations of other relations). Therefore, relational structures between comparable representations can vary in their complexity. Understanding how humans are able to perform relational comparisons of increasing complexity has implications for ontogeny (Halford, Wilson, & Phillips, 1998; Halford, 1992), phylogeny (Premack, 2010; Penn, Holyoak, & Povinelli, 2008), and cognitive neuroscience (Holyoak & Kroger, 1995; Robin & Holyoak, 1995).

Patient and fMRI research across the past two decades has solidified the role of the pFC in integrating relational information of increasing complexity (see Krawczyk, 2012). Frontotemporal dementia patients with selective pFC degeneration (frontal variant) demonstrated pronounced deficits in integrating multiple relationships and inhibiting distracting information compared to those with selective anterior temporal damage (Krawczyk et al., 2008; Morrison et al., 2004; Waltz et al., 1999). When solving nonsemantic Raven's Progressive Matrices (Raven, 1941), the process of integrating increasingly complex relations in the face of distracting elements selectively activated more anterior pFC regions, including the rostrolateral and dorsolateral pFC (Kroger et al., 2002; Christoff et al., 2001; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997). Voxel-based morphometry and symptom lesion mapping techniques have more specifically implicated the left rostrolateral pFC as an essential region for relational integration (Aichelburg et al., 2016; Urbanski et al., 2016) with additional support from both semantic (Green, Kraemer, Fugelsang, Gray, & Dunbar, 2010; Bunge, Wendelken, Badre, & Wagner, 2005) and nonsemantic (Volle, Gilbert, Benoit, & Burgess, 2010; Bunge, Helskog, & Wendelken, 2009) relational mappings.

Inhibiting distracting or irrelevant information from entering working memory is also important for successful relational reasoning. Individuals with greater general fluid intelligence—as measured via Raven's Advanced Progressive Matrices—were more successful in overcoming distracting lure trials during an n-back task and exhibited increased lateral pFC activation during interference control, suggesting relational integration and inhibitory control are closely related and rely on lateral pFC (Gray, Chabris, & Braver, 2003). However, factorially manipulating relational complexity and the number of distracting elements emphasized the role of the lateral frontal pole during relational integration, whereas inhibitory control relied more on lateral pFC regions (i.e., middle and inferior frontal gyri; Cho et al., 2010). Together, these results suggest that integrating increasingly complex relations relies on anterior regions of pFC and competes with cognitive resources in working memory, especially when encoded distracting elements need inhibition for successful relational reasoning (see also Cho, Holyoak, & Cannon, 2007).

Halford et al. (1998) argued that working memory capacity is defined not by the number of items requiring processing but rather by their relational complexity. They defined relational complexity as the "number of related dimensions or sources of variation" (p. 803), such that unary relations have a single relational argument (e.g., fruit[apple]), binary relations have two arguments (e.g., opposite[black, white]), ternary relations have three arguments (e.g., taller[John, Mark, Luke]), and so on. Higher-order relations and relational structures are composed by nesting these first-order relations as arguments in second-order relations. For example, an analogy under this formulation is stated: same [opposite(black, white), opposite(noisy, quiet)]. We borrow from Robin and Holyoak (1995) to characterize higher-order relational complex into three levels:[1]

(1) Attribute/perceptual mappings contain a single dimension same(triangle, triangle)
(2) Relational/analogical mappings contain two dimensions same[same(triangle, triangle), same(square, square)]
(3) System mappings contain three dimensions same(same[same(triangle, triangle), same(square, square)], same[different(circle, rectangle), different (star, diamond)])

Studies of comparative psychology have employed match-to-sample (MTS) tasks to probe the relational abilities of various animal species. MTS tasks present a "sample" stimulus composed of shapes or objects, and participants/subjects are instructed or trained to select a matching "target" stimulus given alternative options. Identical perceptual matches (i.e., matching A to A and not B) are easily performed by vertebrates, including pigeons (e.g., Blaisdell & Cook, 2005), monkeys (e.g., baboons; Bovet & Vauclair, 2001), crows (Smirnova, Zorina, Obozova, & Wasserman, 2015), chimpanzees (Thompson & Oden, 2000; Premack, 1983), humans (Kroger, Holyoak, & Hummel, 2004), and even the invertebrate honeybee (Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001). Importantly, these species have demonstrated the ability to perform same/different discriminations on stimuli not presented during training, also known as transfer, suggesting their ability to acquire a rudimentary conceptual understanding of "sameness" that extends beyond perceptual features. Relational MTS (RMTS) conditions require the selection of targets that are relationally similar to the sample. Analogical matches (i.e., matching AA to BB and not CD) present a critical condition where many animals—other than chimpanzees, crows, and humans—fail to demonstrate relational abilities (Holyoak & Thagard, 1995), such as rhesus monkeys (Flemming, Beran, Thompson, Kleider, & Washburn, 2008). Only after intensive language training using a symbol-based system do chimpanzees demonstrate analogical abilities (Premack, 1983); however, hooded crows have remarkably demonstrated spontaneous analogical transfer during RMTS (Smirnova et al., 2015). Simply put, RMTS tasks have demonstrated that no other animal species, other than humans, can produce complex relational solutions as easily with minimal instruction or training. Whether analogical ability is a uniquely human ability is highly debated (see Premack, 2010; Penn et al., 2008) and is not the focus of this investigation.

Rather, we were interested in the relational abilities of humans that extend beyond analogical matches that Holyoak and Thagard (1995) surmised as exceeding the limit of chimpanzee relational abilities. Comparing the relations between relations, defined above as a system mapping, requires thinking about second-order relations, in all likeliness a unique human ability. Only one study to date has specifically addressed human RMTS performance across the three levels of increasing relational complexity: perceptual, analogical, and system mappings. Kroger et al. (2004) demonstrated that increases in relational complexity increased processing demands as measured via RTs for correctly solved problems. This pattern of results was seen even after controlling for working memory demands and iconic memory effects.

In the interest of isolating processing demands in working memory, Kroger et al. (2004) carefully trained participants on the relational structure in each condition. We developed a novel task that removed this training requirement to observe the unstructured learning rates of RMTS problems with increasing relational complexity and interference. RMTS problem stimuli are composed of fractal-like patterns to reduce the influence of semantic processing. Participants were not instructed on how to solve the problems but received feedback after each attempt to facilitate learning. Because of the increased working memory demands imposed with increasing relational complexity, as well as interference across differing relational structures, we hypothesized that increases in relational complexity would negatively affect participant accuracy. Given that participants received feedback after each trial, we also hypothesized that participants would learn the relational structure of each condition at different rates depending on their relational complexity. Therefore, we predicted that participants would learn the perceptual matches fastest, followed by analogical matches, and then system mappings. In the pursuit of these ideas, two experiments were conducted to further understand factors that contribute to variability, as well as improvements, in participant performance when solving RMTS problems.

EXPERIMENT 1

Methods

Participants

Thirty-nine undergraduate students attending The University of Texas at Dallas participated in the experiment in exchange for course credit. One participant was excluded from analyses because of extremely long RTs (> 10 sec on average) compared with the rest of the sample; therefore, 38 participants were included in the analyses presented below (age: M = 20.70 years, SD = 3.14 years; 14 men; three left-handed; education: M = 14.90 years, SD = 2.17 years). All participants gave informed consent, and experimental procedures were carried out in accordance with the Declaration of Helsinki and approved by the university's institutional review board.

Materials

RMTS task. We developed a novel RMTS task that uniquely features custom-generated fractal-like images presented as stimuli with no practice and limited instruction. Each RMTS problem was composed of nonsemantic fractal-like patterns that were generated using a modified algorithm based on the work of Miyashita, Higuchi, Sakai, and Masui (1991) in MATLAB (R2014b; algorithm and task code is available in an online GitHub repository: https://github.com/mkmiecik14/fractal-rmts). Four hundred ninety-five unique fractal-like images were created such that each problem presented new stimuli that were never repeated throughout the experiment. The fractal-like images were randomly colored across the grayscale spectrum to provide contrast among overlapping elements. To reduce the influence of semantic representations, each problem was visually inspected to ensure stimuli did not resemble shapes seen in everyday life.

The generated fractal-like images were used to create three problem types: perceptual, analogical, and system mapping. Each problem type was constructed using three 2 × 2 configurations of fractal images placed against a black background. A single sample set was placed above two target sets that were separated by a white dotted line resembling the traditional format of an MTS task (see Figure 1A). Matches between sample and target sets were determined using only two relationships of "same" and "different." The three problem types were identical in their stimulus presentation but differed in the relational complexity required to select the correct target.

Correct responses to perceptual mapping problems required participants to regard the sets as whole images and select the target set that contained perceptually identical stimuli (i.e., relationship of sameness) represented in the sample set.

Analogical mapping problems required participants to regard sets as separable top and bottom pairs of fractals that were either the same or different when evaluating the relationship between the top and bottom pairs of fractals within each set. Within-pair fractals, or fractals that appeared on the same row of the 2 × 2 configuration, on analogical mapping problems never differed and were always identical images. Correct analogical maps between sample and target sets required identifying identical relationships shared between top and bottom pairs of fractals (see Figure 1A). These relationships varied between both "same" and "different," thus making it possible to match based on differentness as well as sameness.

System mapping problems, like analogical mapping problems, required pairwise comparisons between top and bottom pairs of stimuli within each set; however, within-pair fractal stimuli now also varied between either being "same" or "different." Relationships now varied within pairs as well

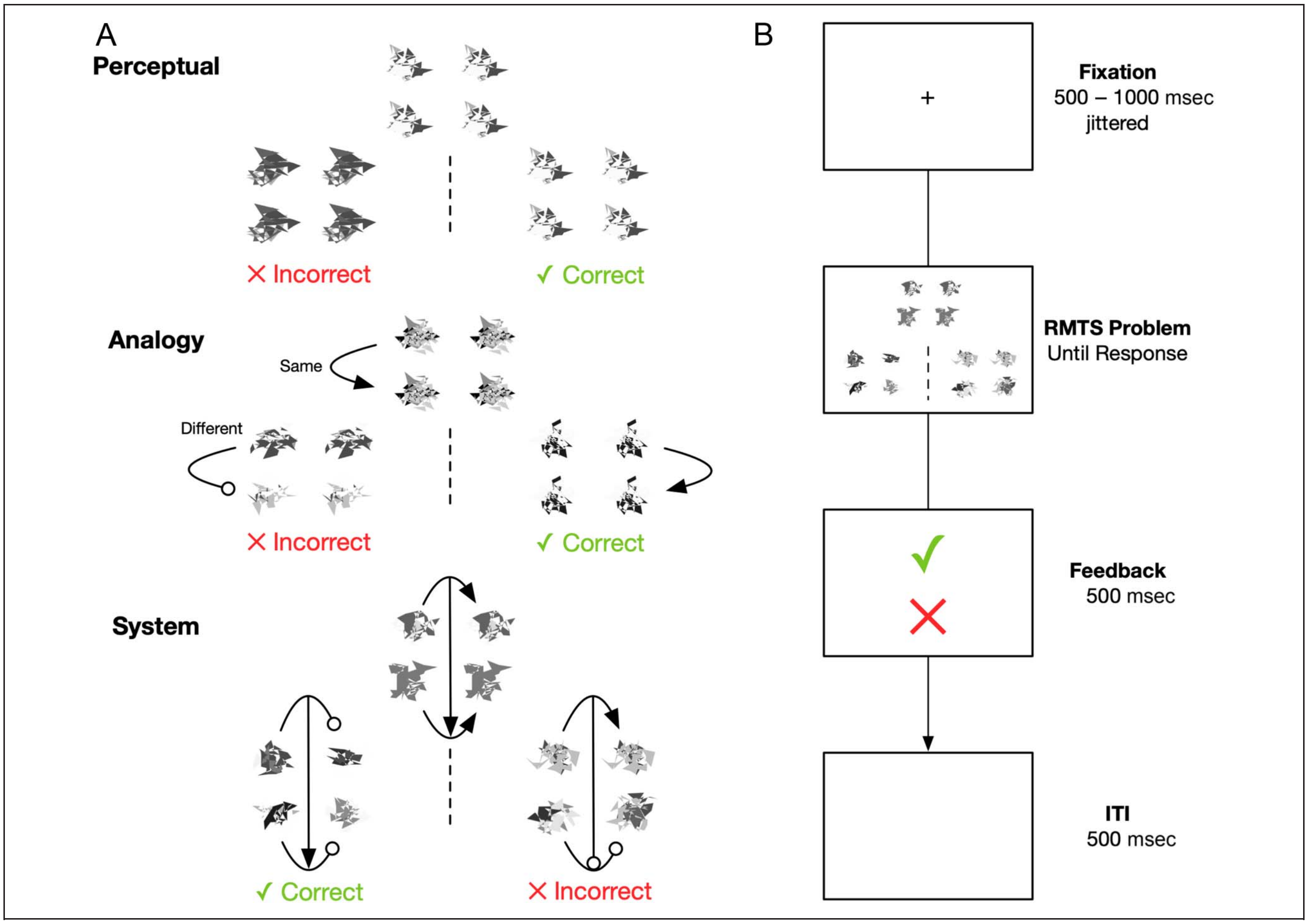

Figure 1. RMTS task conditions and trial procedure. (A) Examples of perceptual, analogical, and system mapping problems (see text for description of counterbalancing and problem types). We intended for participants to infer relationships of same (filled arrows) and different (open circles) that differed based on relational complexity. Correct target choices contained the same relational structure as featured in the sample; therefore, matches were based on same (shown here in analogy and system) and different (not shown). By this definition, perceptual matches were only allowed to match based on same. Correct and incorrect choices are denoted by green checkmarks and red Xs, respectively. (B) Participants received feedback after each of the 96 trials and breaks between blocks every 24 trials. All trials were presented on a black background (white background shown here for illustrative purposes).

as between pairs. Correct responses to system mapping problems required participants to evaluate systems of relationships, evaluating the relationships shared within and between pairs as well as across source and target sets (see Figure 1A).

Thirty-two problems were constructed for each of the three conditions, for a total of 96 problems. All problems were designed such that inference of relational structures other than those outlined above would result in chance performance. In other words, vertical (i.e., column-wise comparisons) or diagonal comparisons within each set would not provide sufficient information to solve the problems above chance. Targets were counterbalanced so that each condition had an equal number of both left- and right-sided correct responses. This resulted in two problem types for perceptual mapping. Analogical mapping problems were represented across four problem types because of equating left and right target response probabilities and counterbalancing between sample–target matching across both relations of "same" and "different." System mapping problems were represented across 32 unique problem types by counterbalancing the following aspects: (1) left and right target response probabilities, (2) equating the probabilities of "same" and "different" relational matching across samples and targets, and (3) ensuring within-pair "same" or "different" relations appeared equally in both top and bottom slots in each of the three sets.

Additional assessments. The RMTS task depicted above was part of a larger study investigating reasoning abilities. Therefore, additional tasks assessing visuospatial working memory, relational abstraction, and analogical reasoning were administered.

Participants' visuospatial working memory was assessed using the Wechsler Memory Scale–Fourth Edition Symbol Span task (Wechsler, 2009). The Symbol Span task required participants to observe a series of nonsemantic shapes for 5 sec and then recall them, in order, on a presented sheet of options. The task progressively increases in difficulty by increasing the number of shapes to memorize within the 5-sec encoding period. Participants' visuospatial working memory ability was quantified using the Wechsler Memory Scale–Fourth Edition Symbol Span total raw score (maximum score possible was 50).

Participants' relational abstraction ability was assessed using the Abstraction and Working Memory (AIM) task (Glahn, Cannon, Gur, Ragland, & Gur, 2000). In this computerized task, participants were shown five shapes on-screen that were distorted slightly to reduce verbalization: two images at the top left, two images at the top right, and one at the center bottom. Participants were instructed to select via button press whether the left or right pair of images best belonged with the bottom image. Five different relational structures based on combinations of color and shape determined correct matches (see Glahn et al., 2000). In the AIM plus Memory subtest, participants were shown the target image for 500 msec that was followed by a blank screen for 2.5 sec before the four shapes appeared at the top. In the AIM simple condition, this additional working memory maintenance requirement was reduced by presenting the target together with the four shapes. Participants' relational abstraction ability was quantified using the number of correct responses out of a possible 20 trials for each condition: (1) AIM and (2) AIM plus memory.

Participants' analogical reasoning ability was assessed using separate verbal and scene-based tasks. The verbal analogies task (Jones, Kmiecik, Irwin, Unsworth, & Morrison, in preparation) presented participants with four-term analogies in the form A:B::C:?, and participants were instructed to select a D term out of two options presented at the bottom of a computer screen. The verbal analogies were manipulated based on semantic distance shared between the source (A:B) and the target (C:D) word pair, thus creating semantically near (JANUARY:MONTH::WINTER:SEASON) and far (LOG:FORT::CERAMIC:JAR) analogies. Furthermore, the options of D and D' were manipulated for distractor salience by varying the association between C and D/D'. Incorrect choices that were highly associated with the C term were more distracting (BOWL:DISH::SPOON:SILVERWARE/FORK) than those with a lower association (OATMEAL:COOKIE::BANANA:MUFFIN/KIWI). We quantified verbal analogy performance as overall task accuracy as percentage correct out of 60 total trials collapsing across semantic distance and distracter salience conditions.

Scene analogy performance was assessed using the Similar Situations Task (SST) developed in-house (Martinez et al., in preparation; Kmiecik, Schauer, Martinez, & Krawczyk, 2016). Briefly, participants were shown 48 line-art scene analogy problems. Each source scene was presented for 5 sec and comprised two sets of items (humans, animals, or objects) that interacted in distinct areas within the scene. One or two arrows directed participants to encode and remember specific items and their relational roles. The target scenes comprised two matching items that interacted analogously to one set of items in the source, whereas two distractor items interacted in a superficially similar manner to the alignable items. No-Match problems were identical to matchable problems but did not contain an analogous match. Participants were tasked with determining which item, if any, was in a similar situation as one of those pointed to in the source by clicking their selection on-screen. Participants were instructed to choose "No Match" if they did not find an analogous match. Therefore, the SST had four conditions: one-arrow match, two-arrow match, one-arrow no match, and two-arrow no match. We quantified scene analogy performance as overall accuracy out of 48 trials by collapsing performance across these conditions.

Procedure

After providing consent, participants were first seated alone in a quiet testing room and administered the SST, verbal analogies, AIM, and Symbol Span tasks. The computerized tasks (i.e., all tasks except for the Symbol Span) were displayed on a 17-in. Dell LCD monitor (resolution: 1280 × 1024). Participants were then administered the RMTS task and were instructed to place their right index and middle fingers on the "1" and "2" keys of the keyboard number pad, respectively. The task was administered using E-Prime 2.0 (SP1; Schneider, Eschman, & Zuccolotto, 2012). The experimenter read the following instructions: "In this experiment, you will be shown three sets of images. One set of images will be shown at the top of the screen, while the other two sets will be shown at the bottom left and bottom right of the screen. Your job is to determine which of the bottom two sets best matches the top set of images." Participants were instructed to press the "1" and "2" keys to choose the bottom left and right images, respectively, and to utilize the feedback after each trial to get as many problems correct as possible.

Each trial began with a randomly jittered fixation cross lasting between 500 and 1000 msec. RMTS problems were presented and remained on-screen until a selection was made. Feedback was presented for 500 msec as a green checkmark for correct responses or a red "X" for incorrect responses (see Figure 1B). A 500-msec intertrial interval separated each trial. All 96 problems were presented in a pseudorandom order across four blocks with 24 problems per block. Each block contained an equal number of left and right correct responses, problems from each condition (i.e., eight), and problem types for perceptual and analogical mapping problems. The 32 unique system mapping problems appeared randomly across the blocks and were never repeated.

Statistical analyses were performed in R (Version 3.5.3; R Core Team, 2015) and RStudio (Version 1.1.456) using the *dplyr* (Wickham, François, Henry, & Müller, 2018), *broom* (Robinson, Hayes, & Couch, 2018), *psych* (Revelle, 2019), and *lmSupport* (Curtin, 2018) packages, and figures were prepared using *ggplot2* (Wickham, 2016) and *patchwork* (Pederson, 2019). Participant performance was modeled using multilevel modeling (see Judd, McClelland, & Ryan, 2017). In Level 1, each participant's performance was modeled individually using a priori orthogonal contrast codes for the within-participant factors of block and condition as well as all their interactions. Because of the four blocks, we modeled three separate block contrasts: linear (−3/2, −1/2, 1/2, 3/2), quadratic (1/2, −1/2, −1/2, 1/2), and cubic (−1/2, 3/2, −3/2, 1/2) changes in performance across time/block. Because of the three different conditions, we modeled two separate condition contrasts: perceptual + analogy versus system mapping (perceptual = 1/3, analogy = 1/3, system = −2/3) and perceptual versus analogy (perceptual = 1/2, analogy = −1/2, system = 0). The contrasts were constructed to first evaluate differences between the highest level of relational complexity (system mappings) and conditions with lower relational complexity (perceptual and analogical matches). Although people solve perceptual and analogical RMTS problems easily (Holyoak & Thagard, 1995), we were further interested in performance differences between these conditions when other relational structures may have interfered with learning rates. Constructing the two contrasts in this manner allowed us to answer questions of theoretical interest while maintaining orthogonality of contrasts that increased the interpretability of the unstandardized regression coefficients (i.e., weighted mean difference). The dependent variables in Level 1 models were proportion correct rates and RTs for correctly solved problems (correct RT). For the correct RT models, the contrast codes stated above were reversed in sign (i.e., multiplied by −1) to reflect predictions that better performance was characterized by reduced RT (i.e., faster responses). The participants' regression estimates (i.e., mean differences between condition contrasts) for each contrast were used as dependent variables in Level 2 models that estimated participant performance in each condition, block, and their interactions (i.e., intercept term).

Results and Discussion

Participant performance on Experiment 1 was well differentiated based on the condition and progress (i.e., block) through the task. The participants learned the perceptual matching problems the fastest with near-ceiling performance throughout the task, whereas analogy and system mapping problems were more slowly learned with variable performance (see Figure 2). Descriptive statistics of condition- and block-wise performance are described in Table 1.

Multilevel modeling results demonstrated several significant within-participant effects for block, condition, and their interactions for both proportion correct and correct RT models (see Table 2). The participants' proportion

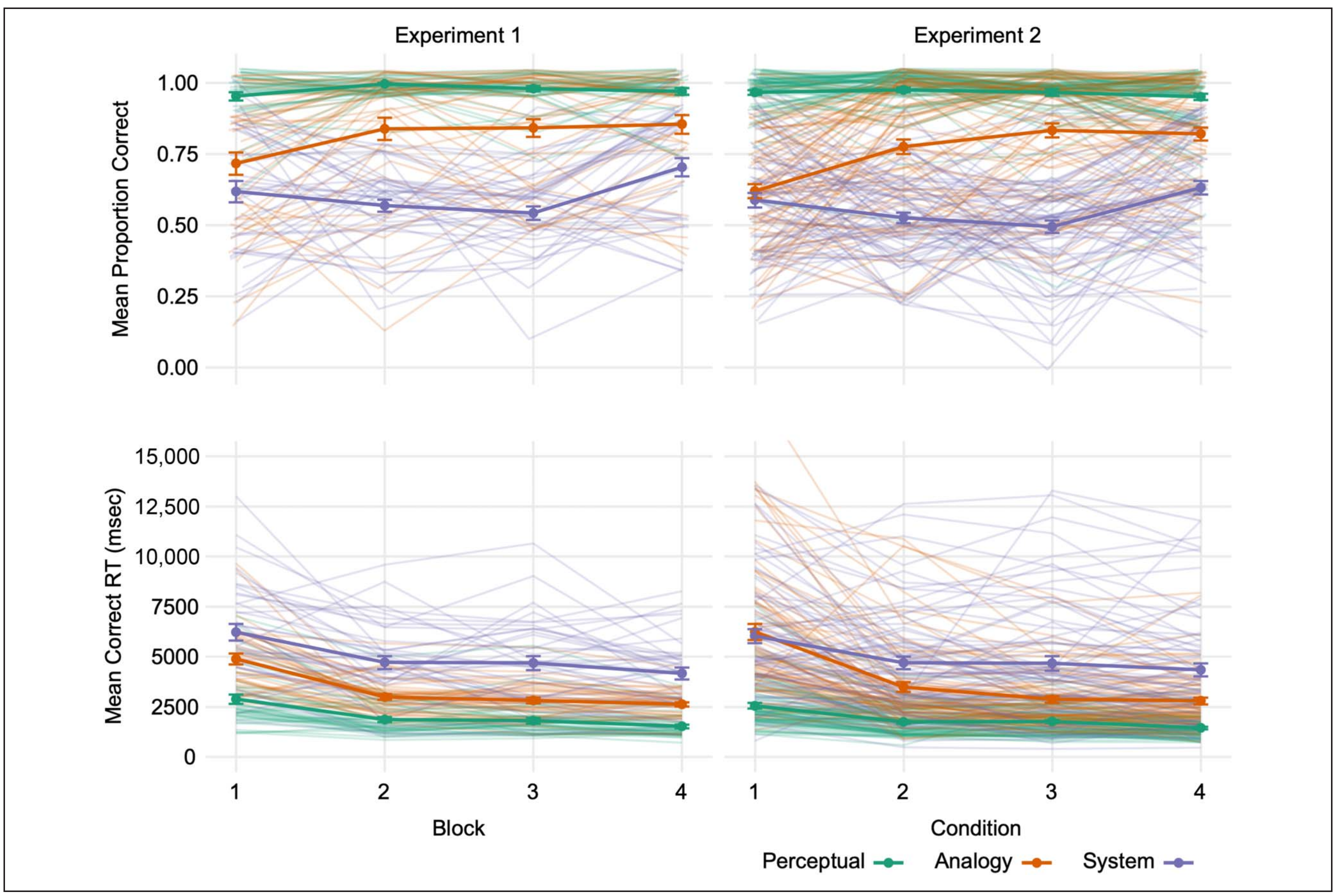

Figure 2. Participant solving performance across block and condition for Experiment 1 (left column) and Experiment 2 (right column) with respect to (top row) proportion correct rates and (bottom row) RT for correct solutions (correct RT). Transparent lines indicate individual participant performance and are slightly jittered to reveal individual differences. Points with *SEM* error bars represent mean performance.

Table 1. RMTS Task Performance between Experiments

Measure	*Condition*	*Experiment*	*Performance across Blocks, M (SD)*			
			1	*2*	*3*	*4*
Accuracy (%)	Perceptual	1	95.39 (8.92)	99.67 (2.03)	98.03 (5.46)	97.04 (7.37)
		2	96.7 (7.27)	97.57 (7.46)	96.7 (10.7)	95.14 (9.74)
	Analogy	1	71.71 (24.26)	83.88 (24.12)	84.21 (19.43)	85.53 (20.24)
		2	61.98 (21.03)	77.6 (22.39)	83.33 (20.56)	82.12 (19.55)
	System	1	61.84 (23.42)	56.91 (13.22)	54.28 (14.61)	70.39 (19.8)
		2	58.85 (21.75)	52.6 (15.55)	49.48 (18.7)	63.19 (20.22)
Correct RT (msec)	Perceptual	1	2890 (1422)	1871 (826)	1815 (852)	1535 (573)
		2	2551 (1199)	1753 (666)	1772 (667)	1455 (529)
	Analogy	1	4894 (1656)	2995 (865)	2823 (871)	2633 (687)
		2	6245 (3420)	3490 (2124)	2883 (1419)	2807 (1428)
	System	1	6230 (2550)	4722 (1981)	4693 (2140)	4171 (1796)
		2	6039 (2974)	4700 (2576)	4678 (3032)	4347 (2755)

One participant was excluded from correct RT analysis in Experiment 2. Correct RT = RT for correct solutions.

correct rates across the four blocks when collapsing across conditions were described by both a linear increase and a fluctuating cubic function, whereas correct RT performance demonstrated linear, quadratic, and cubic changes. When collapsing performance across block, participants were more accurate and faster at solving perceptual and analogy problems (combined) compared to system mapping problems. Furthermore, participants were more accurate and faster at solving perceptual problems compared to analogies.

These mean differences of condition interacted with presentation block. The participants improved in their analogy performance with both linear (proportion correct and correct RT) and quadratic (correct RT only) changes across blocks, whereas perceptual matching performance did not change from high-ceiling performance. In addition, a quadratic effect of block for proportion correct rates was observed that depended on the perceptual and analogy (combined) versus system mapping condition contrast. Experiment 1 results suggest markedly different performance patterns across blocks depending on the relational complexity of the problem. For proportion correct rates, system mapping problems were characterized by a quadratic function, with a decrease before an increase in performance across time, whereas performance for solving analogies tended to linearly increase across the task blocks. Perceptual matching performance remained at ceiling across all four blocks. Correct RTs mainly differentiated perceptual from analogy performance across blocks, demonstrating improvements in learning of analogies via faster RTs with both linear and rapid initial improvement (i.e., quadratic) compared to perceptual matches that were quickly learned.

Taken together, these results support our hypothesis that increasing levels of relational complexity differentially affect participant learning rates for relational structures. Despite minimal instructions, participants immediately learned the perceptual matches. Near-perfect performance on perceptual matchings was expected and further demonstrates the salience that perceptual "sameness" exerts on human decision-making (Holyoak & Thagard, 1995), especially given that each perceptual RMTS trial is test of transfer (i.e., all trials used unique, never repeated, stimuli). Analogy and system mapping performance was characterized by a positive linear slope and a nonlinear quadratic function, respectively, across trial blocks. The difference in these learning rates suggests that increasing relational complexity does not result in monotonic changes to working memory demand; otherwise, participant performance across increasing levels of relational complexity should have resulted in similar linear patterns across blocks.

We believe our experimental design introduced additional interference of relational structures by presenting all conditions equally and randomly within each block. Given that participants received minimal instructions and no practice trials before RMTS task administration, it is likely that the analogical relational structure interfered with participants' ability to learn the system mappings. The initial decrease in system mapping performance is perhaps explained either by the participants' application of the analogical relational structure to system mapping problems or by the application of other incorrect strategies. For example, the application of a vertical or diagonal, rather than the imposed horizontal, solution strategy to any of the three conditions would result in chance levels of performance (see Methods). In addition, incorrectly applying a relational structure from a different condition would also result in chance performance. Therefore, we

Table 2. Experiment 1 Regression Results

Measure	*Source*	*b*	*SS*	*MSE*	*F*	*p*	*PRE*
Proportion correct	Between-participant	0.799	24.263	0.011	2231.398	*	.984
	Linear Block	0.023	0.020	0.001	14.968	*	.288
	Quadratic Block	0.008	0.003	0.006	0.420	.521	.011
	Cubic Block	0.012	0.005	0.001	6.564	.015	.151
	Perceptual + Analogy vs. System	0.286	3.103	0.006	556.048	*	.938
	Perceptual vs. Analogy	0.162	0.997	0.031	31.816	*	.462
	Linear Block × Perceptual + Analogy vs. System	*	*	0.005	0.002	.964	*
	Linear Bock × Perceptual vs. Analogy	−0.038	0.056	0.006	8.742	.005	.191
	Quadratic Block × Perceptual + Analogy vs. System	−0.146	0.805	0.031	25.752	*	.410
	Quadratic Block × Perceptual vs. Analogy	0.028	0.030	0.019	1.551	.221	.040
	Cubic Block × Perceptual + Analogy vs. System	−0.007	0.002	0.005	0.355	.555	.009
	Cubic Block × Perceptual vs. Analogy	−0.006	0.001	0.004	0.411	.525	.011
Correct RT	Between-participant	3439	449,525,947	938,032	479.22	*	.928
	Linear Block	576	12,612,017	156,241	80.72	*	.686
	Quadratic block	−572	12,443,213	603,149	20.63	*	.358
	Cubic block	164	1,015,956	56,565	17.96	*	.327
	Perceptual + Analogy vs. System	2272	196,161,789	1,807,925	108.50	*	.746
	Perceptual vs. Analogy	1308	65,047,834	441,803	147.23	*	.799
	Linear Block × Perceptual + Analogy vs. System	67	168,493	353,708	0.48	.494	.013
	Linear Bock × Perceptual vs. Analogy	283	3,051,728	330,301	9.24	.004	.200
	Quadratic Block × Perceptual + Analogy vs. System	119	536,851	1,487,774	0.36	.552	.010
	Quadratic Block × Perceptual vs. Analogy	−485	8,943,544	1,004,730	8.90	.005	.194
	Cubic Block × Perceptual + Analogy vs. System	51	97,759	320,330	0.31	.584	.008
	Cubic Block × Perceptual vs. Analogy	55	116,768	71,226	1.64	.208	.042

All effects are the intercept term; degrees of freedom for each source and error are 1 and 37, respectively. PRE = proportional reduction in error (η_p^2).

* Values < .001.

surmise that the additional relational interference across conditions contributed to the nonmonotonic differences between levels of increasing relational complexity.

EXPERIMENT 2

The system mapping problems in Experiment 1 were the most relationally complex problems. Solving the system mapping problems required the participants to perform an analogy of analogies. However, we were still surprised at the participants' rather low performance on the system mapping problems in Experiment 1. Participants achieved a mean accuracy of only 70.39% (*SD* = 19.8%) in the final block, despite having 24 learning trials across the first three blocks. Given the minimal instructions presented before the task, it is likely that some participants may have used nonrelational strategies (e.g., perceptual strategies based on color matching or orientation of fractal-like shapes) leading to poor performance and learning. One method thought to foster relational reasoning and invite deeper relational comparisons is to utilize relational language (Gentner, 2016; Vendetti et al., 2015). Gick and Holyoak (1980, 1983) demonstrated that spontaneous analogical transfer between semantically disparate domains is difficult but improves when reasoners are given explicit hints to compare these domains. Although the fractal-like stimuli are minimally semantic, using explicit relational language in the instructions may encourage relational comparisons and discourage more perceptual strategies.

Therefore, to improve system mapping performance, we presented a new sample of participants with the same task as described in Experiment 1 except for a slight instructional manipulation. In Experiment 2, participants were given an additional line of instructions that read "Be sure to attend to how the images are same or different" before beginning the task. We hypothesized that this subtle additional line of instructions would encourage relational thinking, while reducing the use of distracting nonrelational solving strategies. We predicted this hint might be sufficient to raise performance in Experiment 2, especially for relationally complex problems like the analogy and system mappings, in the form of faster learning rates, overall higher accuracies, and faster correct RTs compared to Experiment 1.

Methods

Participants

Seventy-two undergraduate students (age: $M = 21.10$ years, $SD = 3.04$ years; 34 men; seven left-handed, one ambidextrous; education: $M = 14.60$ years, $SD = 1.45$ years) attending The University of Texas at Dallas participated in the experiment in exchange for course credit. One participant was excluded from the correct RT analysis because of incorrectly answering all system mapping problems in Block 3; therefore, the final sample for the proportion correct analysis was $n = 72$, whereas that for the correct RT analysis was $n = 71$. All participants gave informed consent, and experimental procedures were carried out in accordance with the Declaration of Helsinki and approved by the university's institutional review board.

Materials and Procedure

All materials and procedures were identical to those described in Experiment 1 (see above) except for an additional line of instructions presented before the participants beginning the task. The participants were read the same instructions in the same order as described above. After these instructions, the participants were read, "Be sure to attend to how the images are same or different." The task began immediately after the participants indicated they understood these instructions.

The same statistical procedure used to model Experiment 1 performance was replicated when estimating Experiment 2 performance. In addition, we modeled the effect of Experiment 1 (-0.5) versus Experiment 2 ($+0.5$) as a between-participant factor in Level 2 of the multilevel model to examine whether the instructional manipulation affected participant performance. The signs of these contrast codes were reversed (i.e., multiplied by -1) for the correct RT analysis to reflect decreased RT associated with better performance.

Results and Discussion

Similar to Experiment 1, the participants' performance across blocks also depended on the relational complexity of the problems in Experiment 2 (see Figure 2 and Table 1 for descriptive statistics) that was well described with both linear and quadratic changes (see Table 3). In contrast to perceptual matches, which were solved at near ceiling throughout the experiment, analogical comparisons were best characterized by a linear increase and quadratic change in both proportion correct rates and correct RTs, suggesting dramatic initial performance increases between Blocks 1 and 2 (see Perceptual vs. Analogy contrasts in Table 3). In comparison to perceptual and analogical comparisons, the more relationally complex system mapping problems were better characterized by a quadratic function than a linear decrease in proportion correct rates, suggesting participants performed worse before improving at the fourth block. When comparing the proportional reduction in error (also known as η_p^2), these results suggest that system mapping performance was better characterized by a quadratic change over block, whereas analogy performance is better characterized as a linear improvement. The results of Experiment 2 replicate those of Experiment 1 by further demonstrating that participants learn the relational structures of perceptual, analogy, and system mapping problems at different rates when given minimal instruction.

Furthermore, we compared whether participant performance across Experiments 1 and 2 was modulated by the additional line of instructions given in Experiment 2 (i.e., "Be sure to attend to how the images are same or different.") by adding "Experiment" as a between-participant factor in a regression model. Participants across Experiments 1 and 2 did not differ in age, $t(108) = 0.61$, $p = .55$, sex, $\chi^2(1) = 0.71, p = .40$, and years of education, $t(108) = -0.86, p = .39$.

We observed three between-participant interactions of experiment, suggesting reliable differences in participant performance across tasks that is likely attributable to the instructional manipulation (multilevel modeling results are presented in Table 4 for proportion correct and Table 5 for correct RTs). More specifically, the effect of experiment interacted with the linear differences between perceptual versus analogy conditions in both proportion correct rates and correct RTs, such that solving analogies improved at a faster rate (i.e., steeper linear slope) when participants were directed to attend to the relational

Table 3. Experiment 2 Regression Results

Measure	*Source*	*b*	*SS*	*MSE*	*F*	*p*	*PRE*
Proportion correct	Between-participant	0.763	41.887	0.008	5329.753	*	.987
	Linear Block	0.023	0.040	0.002	22.047	*	.237
	Quadratic Block	0.001	*	0.007	0.014	.907	*
	Cubic Block	0.006	0.003	0.001	2.887	.094	.039
	Perceptual + Analogy vs. System	0.304	6.637	0.011	590.282	*	.893
	Perceptual vs. Analogy	0.203	2.958	0.028	105.422	*	.598
	Linear Block × Perceptual + Analogy vs. System	0.020	0.030	0.007	4.332	.041	.058
	Linear Bock × Perceptual vs. Analogy	−0.072	0.370	0.005	67.605	*	.488
	Quadratic Block × Perceptual + Analogy vs. System	−0.148	1.577	0.048	33.160	*	.318
	Quadratic Block × Perceptual vs. Analogy	0.072	0.374	0.025	14.855	*	.173
	Cubic Block × Perceptual + Analogy vs. System	−0.012	0.010	0.007	1.478	.228	.020
	Cubic Block × Perceptual vs. Analogy	−0.002	*	0.004	0.074	.786	.001
Correct RT	Between-participant	3560	899,928,002	2,210,954	407.03	*	.853
	Linear Block	643	29,345,783	233,422	125.72	*	.642
	Quadratic Block	−695	34,267,984	716,922	47.80	*	.406
	Cubic Block	147	1,524,606	75,138	20.29	*	.225
	Perceptual + Analogy vs. System	2072	304,678,497	2,823,922	107.89	*	.607
	Perceptual vs. Analogy	1973	276,523,546	2,017,098	137.09	*	.662
	Linear Block × Perceptual + Analogy vs. System	−200	2,830,433	778,895	3.63	.061	.049
	Linear Bock × Perceptual vs. Analogy	765	41,561,072	836,341	49.69	*	.415
	Quadratic Block × Perceptual + Analogy vs. System	286	5,819,810	3,719,954	1.56	.215	.022
	Quadratic Block × Perceptual vs. Analogy	−1099	85,766,297	2,708,282	31.67	*	.311
	Cubic Block × Perceptual + Analogy vs. System	24	40,787	376,437	0.11	.743	.002
	Cubic Block × Perceptual vs. Analogy	46	152,896	191,734	0.80	.375	.011

All effects are the intercept term; degrees of freedom for each source and error are 1 and 71 (proportion correct) and 1 and 70 (correct RT), respectively.

* Values < .001.

properties of the problems (Experiment 2) compared to when these additional instructions were not given (Experiment 1; see Figure 3). We also observed a quadratic difference between perceptual and analogy correct RTs, suggesting that the relational hint impacted participant solving strategies resulting in facilitated initial acquisition of analogies compared to perceptual matches. Together, improved proportion correct rates and faster correct RTs suggest that our instructional manipulation may have facilitated relational comparisons by directing participants' attention to relationally compare items. However, this interpretation is limited given the between-participant nature of this design (e.g., Experiment 2 participants performed worse in Block 1 so they had more opportunity to improve). Our primary aim was to examine learning rates of complex relational structures over time with

Table 4. Regression Results for Proportion Correct Rates between Experiments

Source	*Term*	*b*	*SS*	*MSE*	*F*	*p*	*PRE*
Between-participant	Intercept	0.781	60.670	0.009	6823.138	*	.984
Between-participant	Experiment	−0.036	0.033	0.009	3.693	.057	.033
Linear Block	Intercept	0.023	0.053	0.002	32.495	*	.231
Linear Block	Experiment	0.001	*	0.002	0.010	.922	*
Quadratic Block	Intercept	0.005	0.002	0.007	0.324	.570	.003
Quadratic Block	Experiment	−0.007	0.001	0.007	0.184	.669	.002
Cubic Block	Intercept	0.009	0.008	0.001	9.279	.003	.079
Cubic Block	Experiment	−0.006	0.001	0.001	1.065	.304	.010
Perceptual + Analogy vs. System	Intercept	0.295	8.640	0.009	928.692	*	.896
Perceptual + Analogy vs. System	Experiment	0.018	0.008	0.009	0.850	.359	.008
Perceptual vs. Analogy	Intercept	0.182	3.308	0.029	113.350	*	.512
Perceptual vs. Analogy	Experiment	0.041	0.041	0.029	1.411	.238	.013
Linear Block × Perceptual + Analogy vs. System	Intercept	0.010	0.010	0.006	1.612	.207	.015
Linear Block × Perceptual + Analogy vs. System	Experiment	0.021	0.011	0.006	1.776	.185	.016
Linear Bock × Perceptual vs. Analogy	Intercept	−0.055	0.302	0.006	52.018	*	.325
Linear Bock × Perceptual vs. Analogy	Experiment	−0.033	0.027	0.006	4.727	.032	.042
Quadratic Block × Perceptual + Analogy vs. System	Intercept	−0.147	2.144	0.042	51.061	*	.321
Quadratic Block × Perceptual + Analogy vs. System	Experiment	−0.002	*	0.042	0.004	.953	*
Quadratic Block × Perceptual vs. Analogy	Intercept	0.050	0.249	0.023	10.769	.001	.091
Quadratic Block × Perceptual vs. Analogy	Experiment	0.044	0.048	0.023	2.093	.151	.019
Cubic Block × Perceptual + Analogy vs. System	Intercept	−0.009	0.008	0.006	1.397	.240	.013
Cubic Block × Perceptual + Analogy vs. System	Experiment	−0.005	0.001	0.006	0.101	.751	.001
Cubic Block × Perceptual vs. Analogy	Intercept	−0.004	0.002	0.004	0.464	.497	.004
Cubic Block × Perceptual vs. Analogy	Experiment	0.004	*	0.004	0.131	.718	.001

Degrees of freedom for each source and error are 1 and 108, respectively.

* $p < .001$.

minimal instruction; therefore, we would argue that a within-participant design would conflict with this aim such that repeated exposure to the task to examine instructional manipulations would confound with previous experience with the task. Although limited in inference, a between-participant design was a necessary concession to examine the navigation of relational structures with limited prior experience and instruction.

When collapsing across experiments, the participants demonstrated several repeated-measures interactions as indicated by significant intercepts. Linear and quadratic learning rates across blocks well described the differences between analogy and perceptual matches for both proportion correct rates and correct RTs. System mapping proportion correct rates, but not correct RTs, were again described by quadratic changes—first decrease before increase in performance—compared to the combined perceptual and analogy performance. These results further suggest that relational complexity interacts with learning rates, however, not in a monotonic pattern. System mappings, which are the most relationally complex, were better characterized by initial difficulties in learning, whereas

Table 5. Regression Results for Correct RTs between Experiments

Source	*Term*	*b*	*SS*	*MSE*	*F*	*p*	*PRE*
Between-participant	Intercept	3500	1,212,732,614	1,770,785	684.86	*	.865
Between-participant	Experiment	−121	361,079	1,770,785	0.20	.652	.002
Linear Block	Intercept	610	36,781,173	206,733	177.92	*	.624
Linear Block	Experiment	−67	110,440	206,733	0.53	.466	.005
Quadratic Block	Intercept	−633	39,732,299	677,580	58.64	*	.354
Quadratic Block	Experiment	122	371,399	677,580	0.55	.461	.005
Cubic Block	Intercept	155	2,379,436	68,716	34.63	*	.244
Cubic Block	Experiment	17	7130	68,716	0.10	.748	.001
Perceptual + Analogy vs. System	Intercept	2172	466,991,510	2,472,595	188.87	*	.638
Perceptual + Analogy vs. System	Experiment	201	995,095	2,472,595	0.40	.527	.004
Perceptual vs. analogy	Intercept	1641	266,595,707	1,472,369	181.07	*	.629
Perceptual vs. analogy	Experiment	−665	10,950,917	1,472,369	7.44	.007	.065
Linear Block × Perceptual + Analogy vs. System	Intercept	−67	438,333	631,868	0.69	.407	.006
Linear Block × Perceptual + Analogy vs. System	Experiment	266	1,754,686	631,868	2.78	.099	.025
Linear Bock × Perceptual vs. Analogy	Intercept	524	27,210,489	661,355	41.14	*	.278
Linear Bock × Perceptual vs. Analogy	Experiment	−482	5,743,519	661,355	8.68	.004	.075
Quadratic Block × Perceptual + Analogy vs. System	Intercept	203	4,063,250	2,948,078	1.38	.243	.013
Quadratic Block × Perceptual + Analogy vs. System	Experiment	−167	693,982	2,948,078	0.24	.629	.002
Quadratic Block × Perceptual vs. Analogy	Intercept	−792	62,121,759	2,119,203	29.31	*	.215
Quadratic Block × Perceptual vs. Analogy	Experiment	614	9,329,818	2,119,203	4.40	.038	.040
Cubic Block × Perceptual + Analogy vs. System	Intercept	37	138,079	357,036	0.39	.535	.004
Cubic Block × Perceptual + Analogy vs. System	Experiment	27	17,715	357,036	0.05	.824	*
Cubic Block × Perceptual vs. Analogy	Intercept	51	256,709	150,063	1.71	.194	.016
Cubic Block × Perceptual vs. Analogy	Experiment	9	2017	150,063	0.01	.908	*

Units for *b*, *SS*, and *MSE* are in milliseconds; degrees of freedom for each source and error are 1 and 107, respectively.

* Values < .001.

analogies were characterized by both gradual (linear) and rapid initial (quadratic) increases in performance.

PERFORMANCE TYPES

When further examining the raw data across the different experiments (see Figure 2), we noticed that several participants achieved rather high proportion correct rates in the final block on system mapping problems than the group mean suggested. The pattern of participant performance suggested two groups of individuals: (1) those who learned the system mappings and (2) those who failed to learn their relational structure. To further examine this idea, we divided participants into two groups based on their system mapping performance on the fourth block. Those participants who correctly answered a minimum of five Block 4 system mapping trials (i.e., above-chance performance) were labeled as higher

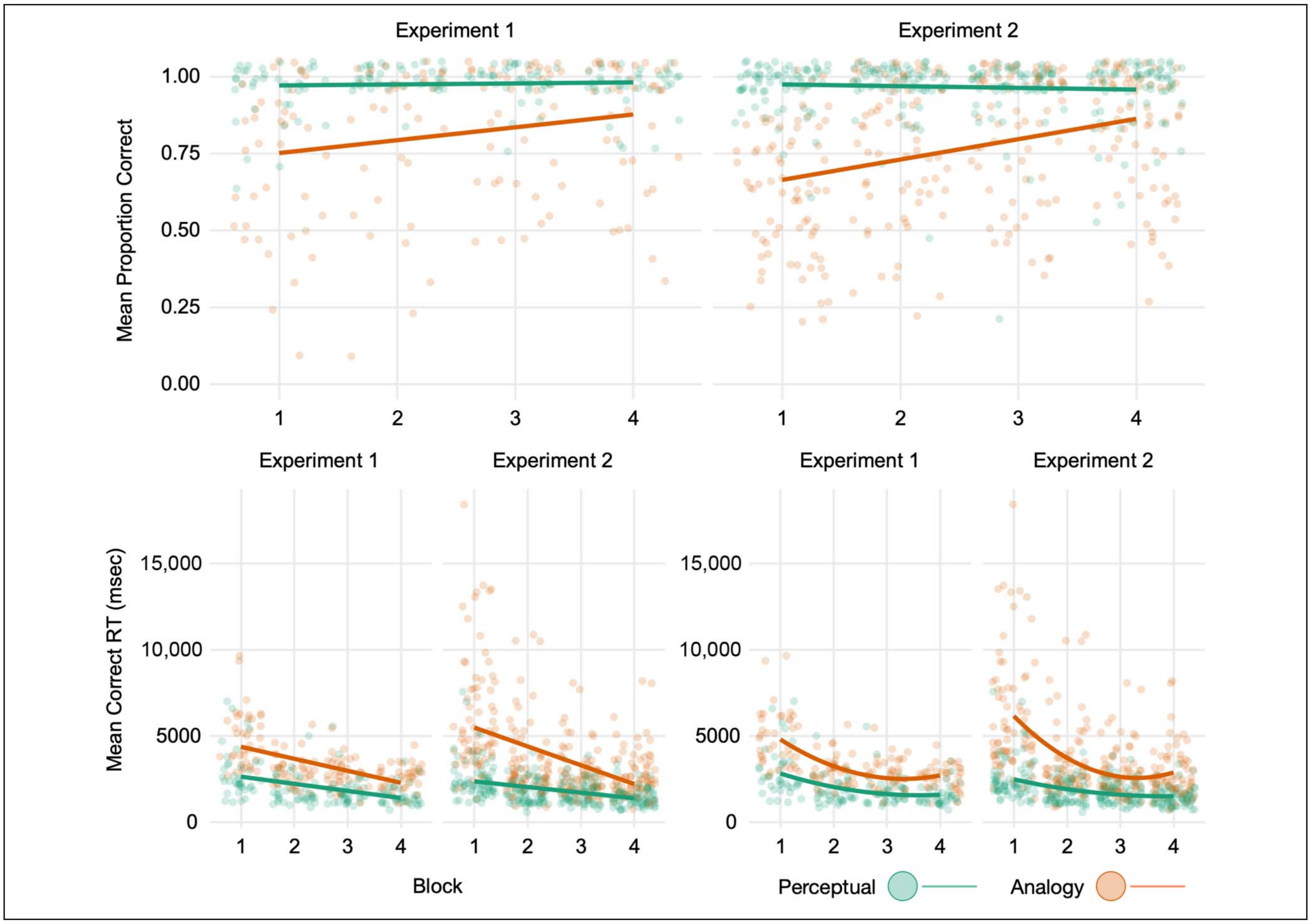

Figure 3. Effect of experiment on (top row) proportion correct rates and (bottom row) RT for correct solutions (correct RT). Before the task, Experiment 2 presented an additional line of instructions cueing participants to pay attention to how stimuli were same or different. Participants in Experiment 2 demonstrated elevated performance as evidenced by steeper linear slopes in the analogy condition for both outcome measures. An additional quadratic fit for correct RTs suggests this initial acquisition for analogies was more rapid in Experiment 2.

performers; otherwise, participants at chance or lower performance were characterized as lower performers. The same statistical procedure used to model the above Experiment 1 versus Experiment 2 differences was replicated here by replacing the "Experiment" factor with a between-participant "Performance type" factor (lower performers = −0.5, higher performers = +0.5). Again, the signs of these contrast codes were reversed (i.e., multiplied by −1) for the correct RT analysis to reflect decreased RT associated with better performance.

Methods

Participants

From the 110 participants, 60 were classified as lower performers (age: M = 21.20 years, SD = 3.47 years; 24 men; six left-handed, one ambidextrous; education: M = 14.80 years, SD = 1.57 years) and 50 were classified as higher performers (age: M = 20.60 years, SD = 2.49 years; 24 men; four left-handed; education: M = 14.60 years, SD = 1.91 years). These two groups did not differ in age, $t(108) = 1.05, p = .30$, sex, $\chi^2(1) = 0.42, p = .52$, and years of education, $t(108) = 0.46, p = .65$. The participant who was excluded from the correct RT analysis in Experiment 2 was a lower performer; therefore, the correct RT analyses compared 59 lower performers to 50 higher performers, whereas proportion correct analyses compared 60 lower performers to 50 higher performers. Because of a computer malfunction, a small subset of participant data was lost for the AIM and AIM + memory (n = 3 higher performers; n = 5 lower performers), verbal analogy (n = 1 lower performer), and scene analogy (n = 1 higher performer) assessments (see Methods under Experiment 1 for task descriptions).

Results and Discussion

Descriptive statistics of task performance are presented in Table 6 for performance types. Multilevel modeling results demonstrated several performance type interactions for proportion correct rates (see Table 7) and correct RTs (see Table 8). Although performance type was only

Table 6. RMTS Task Performance between Higher and Lower Performers

Measure	*Condition*	*Performer*	*Proportion Correct across Blocks, M (SD), %* 1	2	3	4
Accuracy (%)	Perceptual	Lower	95.42 (8.6)	97.71 (7.8)	95 (12)	93.75 (10.92)
		Higher	97.25 (6.82)	99 (3.43)	99.75 (1.77)	98.25 (5.06)
	Analogy	Lower	57.29 (20.1)	72.08 (22.83)	75.42 (21.21)	75.83 (21.57)
		Higher	75 (21.72)	89 (19.99)	93.5 (13.18)	92.25 (12.6)
	System	Lower	52.92 (18.75)	50.63 (16.02)	46.04 (16.19)	49.79 (12.71)
		Higher	68.25 (23.45)	58.25 (12.27)	57.25 (17.14)	84.75 (6.33)
Correct RT (msec)	Perceptual	Lower	2345 (1061)	1685 (668)	1663 (696)	1377 (505)
		Higher	3053 (1424)	1923 (772)	1934 (755)	1609 (565)
	Analogy	Lower	5279 (2801)	3275 (1986)	2709 (1337)	2584 (1311)
		Higher	6359 (3125)	3367 (1569)	3043 (1129)	2939 (1087)
	System	Lower	5221 (2695)	4135 (2397)	3505 (2346)	3475 (2418)
		Higher	7150 (2629)	5384 (2185)	6074 (2536)	5243 (2155)

One lower performer was excluded from correct RT analysis.

determined by system mapping accuracy on the fourth block, this factor differentiated performance on the entire task such that higher performers, despite responding 902 msec slower, correctly solved 13% more RTMS problems than lower performers. Furthermore, higher and lower performers were differentiated in RMTS analogy performance. Perceptual matching performance was comparable between the groups, although higher performers correctly answered more analogy problems than lower performers (see Figure 4). In contrast, perceptual and analogy conditions did not differ between performance types for correct RTs; rather, correct RTs differentiated system mappings from analogy and perceptual matches (combined) such that higher performers took longer to solve system mappings compared to lower performers (see Figure 4). We interpret these effects to mean that higher performers were generally more relationally minded than lower performers and were better able to discern the difference between analogy and system mapping relational structures despite interference; however, this increase in accuracy for higher performers resulted in increased RTs that likely reflect a speed–accuracy tradeoff driven by the relational comparisons.

Importantly, block and condition effects interacted, resulting in performance type interactions in proportion correct rates with linear and quadratic changes in perceptual + analogy (combined) versus system mapping performance (see Figure 4). When comparing the effect of performance type to the effect of experiment (i.e., instructional manipulation), as presented above, it is clear that performance types were stronger in differentiating performance than the instructional manipulation presented in Experiment 2. Higher performers—those who achieved above-chance performance on system mapping problems by the end of the experiment—learned the various relational structures at a different rate than lower performers. This ability to learn the system mapping relational structure also benefited in the analogy condition, with higher performers achieving greater solving rates than lower performers.

When inspecting correct RTs, performance type interacted with the cubic effect of block on perceptual + analogy (combined) versus system mappings (see Figure 4). While solving system mappings, higher performers responded slower in Block 3 after an initial improvement with faster correct RTs between Blocks 1 and 2. In contrast, lower performers did not perform slower for correct problems during Block 3 system mappings but responded increasingly faster throughout the course of the experiment. Combined with the low performance in Block 3 system mappings (inaccurate and slow RTs), higher performers appear affected by interference of competing relational structures, likely the analogy structures. To further explore this possibility, we computed Pearson correlations with bootstrapped 95% confidence intervals (CIs) between analogy and system mapping performance across block and performance type for proportion correct rates and correct RTs. The presence of negative correlations would support the idea that increases in analogy solving would interfere and result in decreased system mapping performance.

The correlations of higher performers for proportion correct rates demonstrated all positive correlations (see Figure 4): Block 1, $r = .52$, 95% CI [.28, .70]; Block 2, $r = .18$ [−.10, .44]; Block 3, $r = .5$ [.25, .68]; and Block 4, $r = .37$ [.10, .59]. Meanwhile, lower performers demonstrated both

Table 7. Regression Results for Proportion Correct Rates between Performance Types

Source	*Term*	*b*	*SS*	*MSE*	*F*	*p*	*PRE*
Between-participant	Intercept	0.781	66.539	0.005	12753.797	*	.992
Between-participant	Performance	0.126	0.430	0.005	82.361	*	.433
Linear Block	Intercept	0.024	0.064	0.001	43.196	*	.286
Linear Block	Performance	0.024	0.015	0.001	10.305	.002	.087
Quadratic Block	Intercept	0.006	0.004	0.006	0.632	.429	.006
Quadratic Block	Performance	0.051	0.072	0.006	11.859	.001	.099
Cubic Block	Intercept	0.008	0.007	0.001	7.968	.006	.069
Cubic Block	Performance	−0.001	*	0.001	0.046	.830	*
Perceptual + Analogy vs. System	Intercept	0.294	9.443	0.008	1165.029	*	.915
Perceptual + Analogy vs. System	Performance	−0.071	0.137	0.008	16.931	*	.136
Perceptual vs. Analogy	Intercept	0.182	3.621	0.024	147.892	*	.578
Perceptual vs. Analogy	Performance	−0.142	0.549	0.024	22.421	*	.172
Linear Block × Perceptual + Analogy vs. System	Intercept	0.011	0.012	0.005	2.261	.136	.021
Linear Block × Perceptual + Analogy vs. System	Performance	−0.058	0.092	0.005	17.164	*	.137
Linear Bock × Perceptual vs. Analogy	Intercept	−0.060	0.387	0.006	64.453	*	.374
Linear Bock × Perceptual vs. Analogy	Performance	0.014	0.005	0.006	0.911	.342	.008
Quadratic Block × Perceptual + Analogy vs. System	Intercept	−0.154	2.600	0.036	72.995	*	.403
Quadratic Block × Perceptual + Analogy vs. System	Performance	−0.159	0.687	0.036	19.298	*	.152
Quadratic Block × Perceptual vs. Analogy	Intercept	0.057	0.355	0.024	15.100	*	.123
Quadratic Block × Perceptual vs. Analogy	Performance	0.006	0.001	0.024	0.039	.843	*
Cubic Block × Perceptual + Analogy vs. System	Intercept	−0.011	0.012	0.006	2.071	.153	.019
Cubic Block × Perceptual + Analogy vs. System	Performance	−0.015	0.006	0.006	1.037	.311	.010
Cubic Block × Perceptual vs. Analogy	Intercept	−0.004	0.001	0.004	0.383	.537	.004
Cubic Block × Perceptual vs. Analogy	Performance	−0.003	*	0.004	0.065	.799	.001

Degrees of freedom for each source and error are 1 and 108, respectively.

* Values < .001.

positive and negative correlations: Block 1, $r = .43$ [.19, .61]; Block 2, $r = -.24$ [−.47, .01]; Block 3, $r = -.06$ [−.31, .20]; and Block 4, $r = .31$ [.06, .52]. Despite the negative correlations not reaching significance (i.e., 95% CIs crossed zero) for lower performers, we observed a trending negative correlation for Block 2 ($p = .06$) and characteristically different patterns in correlations between higher and lower performers, especially in Block 3, providing supportive evidence of interference in lower performers. In contrast, analogy and system mapping correct RTs were all positively correlated and significant for higher performers —Block 1, $r = .46$ [.20, .65]; Block 2, $r = .47$ [.22, .66]; Block 3, $r = .58$ [.36, .74]; and Block 4, $r = .62$ [.41, .77] —and lower performers—Block 1, $r = .52$ [.31, .69]; Block 2, $r = .49$ [.27, .67]; Block 3, $r = .68$ [.52, .80]; and Block 4, $r = .77$ [.64, .86]. These similar positive correlations suggest that RTs, when correctly solving analogy and system mapping RMTS problems, shared similar solving patterns between higher and lower performers; however, correct RTs speak less toward interference effects because of their inclusion of only correctly solved trials and therefore, by definition, have fewer trials.

Table 8. Regression Results for Correct RTs between Performance Types

Source	*Term*	*b*	*SS*	*MSE*	*F*	*p*	*PRE*
Between-participant	Intercept	3555	1,368,415,062	1,568,305	872.54	*	.891
Between-participant	Performance	−902	22,026,419	1,568,305	14.04	*	.116
Linear Block	Intercept	623	42,022,828	206,025	203.97	*	.656
Linear Block	Performance	−83	186,285	206,025	0.90	.344	.008
Quadratic Block	Intercept	−661	47,315,312	668,803	70.75	*	.398
Quadratic Block	Performance	220	1,310,519	668,803	1.96	.164	.018
Cubic Block	Intercept	161	2,803,064	58,168	48.19	*	.311
Cubic Block	Performance	−205	1,135,748	58,168	19.53	*	.154
Perceptual + Analogy vs. System	Intercept	2202	524,876,801	1,938,997	270.70	*	.717
Perceptual + Analogy vs. System	Performance	−1465	58,090,135	1,938,997	29.96	*	.219
Perceptual vs. analogy	Intercept	1746	329,974,807	1,572,021	209.90	*	.662
Perceptual vs. analogy	Performance	−103	288,182	1,572,021	0.18	.669	.002
Linear Block × Perceptual + Analogy vs. System	Intercept	−117	1,486,202	632,439	2.35	.128	.021
Linear Block × Perceptual + Analogy vs. System	Performance	250	1,693,559	632,439	2.68	.105	.024
Linear Bock × Perceptual vs. Analogy	Intercept	599	38,891,035	714,306	54.45	*	.337
Linear Bock × Perceptual vs. Analogy	Performance	−54	77,816	714,306	0.11	.742	.001
Quadratic Block × Perceptual + Analogy vs. System	Intercept	245	6,512,447	2,909,933	2.24	.138	.020
Quadratic Block × Perceptual + Analogy vs. System	Performance	−420	4,775,496	2,909,933	1.64	.203	.015
Quadratic Block × Perceptual vs. Analogy	Intercept	−897	87,096,970	2,185,321	39.86	*	.271
Quadratic Block × Perceptual vs. Analogy	Performance	289	2,255,201	2,185,321	1.03	.312	.010
Cubic Block × Perceptual + Analogy vs. System	Intercept	46	230,144	332,838	0.69	.408	.006
Cubic Block × Perceptual + Analogy vs. System	Performance	−310	2,606,838	332,838	7.83	.006	.068
Cubic Block × Perceptual vs. Analogy	Intercept	53	306,006	148,144	2.07	.154	.019
Cubic Block × Perceptual vs. Analogy	Performance	−88	207,351	148,144	1.40	.239	.013

Units for *b*, *SS*, and *MSE* are in milliseconds; degrees of freedom for each source and error are 1 and 107, respectively.

* $p < .001$.

Individuals classified as higher performers above also performed better on separate tasks of visuospatial working memory (i.e., symbol span), abstraction (i.e., AIM and AIM + memory), and verbal and scene-based analogical reasoning (i.e., verbal analogies and the SST, respectively; see Table 9). This widespread differentiation in performance types suggests that the RMTS task used in this investigation provided an excellent measure of reasoning that was sensitive to individual differences and general fluid intelligence abilities. Higher performers demonstrated competence and understanding of system mappings given their elevated Block 4 performance (M = 85%, SD = 6%) compared to lower performers who hovered around chance performance (M = 50%, SD = 13%). Given higher performers additionally performed better across all additional assessments that required attention to and maintenance of relational information, even across temporal delays in the case of AIM + memory and SST, it is likely that these individuals have elevated working memory capacities to (1) process relationally complex situations that are novel and (2) face interference. Although hardly comprehensive, our cognitive battery contained

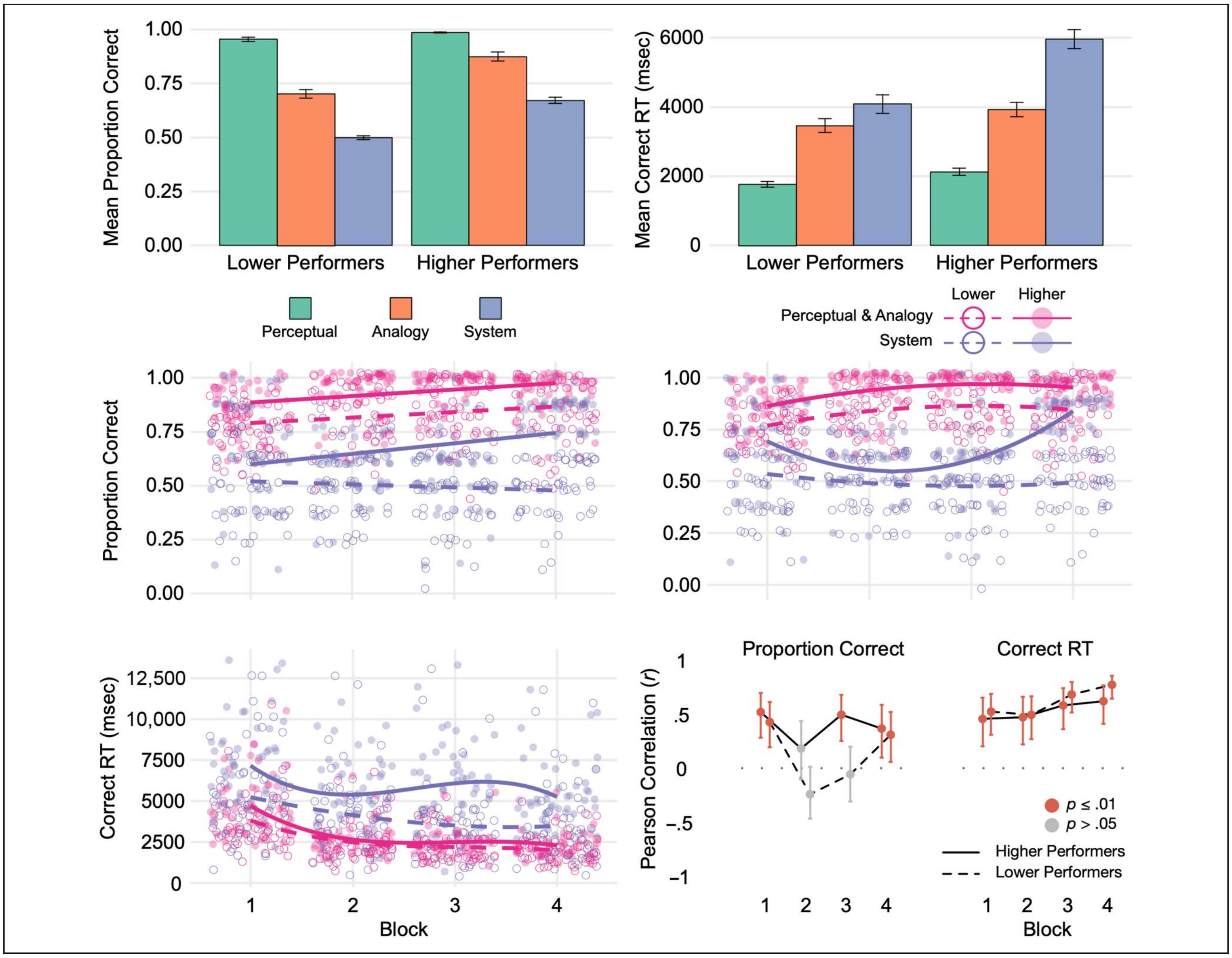

Figure 4. Effect of participant performance types on overall performance and learning rates (data shown are collapsed across both experiments). Higher performers achieved above-chance performance during Block 4 system mappings. Top row depicts condition-wise mean performance for (left) proportion correct rates and (right) RTs for correct solutions (correct RT) with *SEM* error bars. (Center row) Block-wise proportion correct rates were differentiated both linearly (left) and quadratically (right) between performance types when comparing perceptual and analogy conditions (combined) to system mappings. Points indicate individual participant performance and are slightly jittered to reveal individual differences, whereas lines indicate predicted regression fits. (Bottom left) Correct RTs for higher performers continuously fluctuated throughout the experiment (cubic fit) with slower RTs in Block 3 despite poor performance. (Bottom right) Correlations with bootstrapped 95% CIs between analogy and system mapping performance suggest that lower performers likely encountered interference from competing relational structures with a trending negative correlation in Block 2 for proportion correct rates (p = .06). Correct RT correlations were all positive and not different between performance types.

Table 9. Additional Assessment Differences between Performance Types

	Performance Type				
Measure	*Lower, M (SD)*	*Higher, M (SD)*	*t*	*df*	*p*
Symbol span	26 (7)	29 (7)	2.05	108	.04
AIM	17 (2)	18 (2)	2.11	100	.037
AIM + memory	17 (2)	18 (1)	3.14	100	.002
Verbal analogies	87% (7%)	92% (8%)	2.85	107	.005
Scene analogies (SST)	57% (19%)	69% (18%)	3.26	107	.001

Differences in degrees of freedom across assessments resulted from computer malfunction-related data loss (see Methods under Performance Types).

several challenging tasks that required participants to face novel situations/stimuli to detect relational patterns and develop strategies for accurate performance.

Conclusion

This investigation explored how humans reason through increasingly complex relationships. Using a novel variant of the RMTS task, we demonstrated that increasingly complex relational structures differentially affected accuracy rates and correct RTs when participants were given problems without practice, minimal instructions, and feedback after each selection. Participants quickly learned perceptual and analogical mappings; however, not all participants learned the most relationally complex structure of system mappings. To improve task performance, we directed the participants' attention to the relational properties of "sameness" and "difference" in Experiment 2. This had a modest effect on improving analogy performance but did not translate into improving system mapping performance. The effect of instructions notwithstanding, whether participants performed above chance on system mappings in the experiments' final block highly differentiated participant performance on not only the RMTS task but also tasks of visuospatial working memory, abstraction, and verbal/scene-based analogy. Given the differentiating linear, quadratic, and even cubic effects on the analogy and system mapping conditions, respectively, we argue that discerning increasingly complex relational structures places nonmonotonic demands on working memory. We believe this is likely a function of our experimental design with participants, especially those who have lower working memory abilities and place less emphasis on relational properties (i.e., lower performers), needing to resolve interference because of differences in relational structures; however, our experiments were not designed to fully test theories of interference in working memory. Therefore, future studies would be well served to further explore how interference competes with relational complexity in working memory.

Our two-experiment study demonstrated that performance on a nonsemantic RMTS task is quite variable. Even young college-educated people do not always effectively solve the system mapping problems, despite when provided immediate feedback on their performance. This ability to learn and understand complex relational structures relies on a variety of cognitive abilities that support relational thinking, including working memory and interference control. System mapping performance remains a dividing line among primates (Penn et al., 2008; Holyoak & Thagard, 1995), and further understanding the cognitive substrates of this divide, as well as the individual differences within humans, will aid our understanding on the basis of relational comparison and related cognitive abilities that contribute to intellect.

Acknowledgments

We thank David Martinez for his assistance in developing the study protocol; Ekarin Pongpipat for his assistance with statistical procedures; and Mina Kim, Garrett Virgin, Alex Martin, Roberto Espinoza, Kimberly Brennan, Pranali Kamat, Brandon Pires, and Niki Allahyari for their help with data collection.

Reprint requests should be sent to Matthew J. Kmiecik, Department of Obstetrics and Gynecology, NorthShore University HealthSystem, 2650 Ridge Ave., Suite 1507, Evanston, IL 60201, or via e-mail: mkmiecik@uchicago.edu.

Note

1. We recognize that there are an infinite number of levels to higher-order relational complexity, but for simplicity and cohesiveness with the current investigation, we limit ourselves to these three levels: perceptual, analogy, and system mappings.

REFERENCES

Aichelburg, C., Urbanski, M., de Schotten, M. T., Humbert, F., Levy, R., & Volle, E. (2016). Morphometry of left frontal and temporal poles predicts analogical reasoning abilities. *Cerebral Cortex*, *26*, 915–932. **DOI:** https://doi.org/10.1093/cercor/bhu254, **PMID:** 25331605

Blaisdell, A. P., & Cook, R. G. (2005). Two-item *same–different* concept learning in pigeons. *Animal Learning & Behavior*, *33*, 67–77. **DOI:** https://doi.org/10.3758/BF03196051, **PMID:** 15971494

Bovet, D., & Vauclair, J. (2001). Judgment of conceptual identity in monkeys. *Psychonomic Bulletin & Review*, *8*, 470–475. **DOI:** https://doi.org/10.3758/BF03196181, **PMID:** 11700897

Bunge, S. A., Helskog, E. H., & Wendelken, C. (2009). Left, but not right, rostrolateral prefrontal cortex meets a stringent test of the relational integration hypothesis. *Neuroimage*, *46*, 338–342. **DOI:** https://doi.org/10.1016/j.neuroimage.2009.01.064, **PMID:** 19457362, **PMCID:** PMC2864011

Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex*, *15*, 239–249. **DOI:** https://doi.org/10.1093/cercor/bhh126, **PMID:** 15238433

Cho, S., Holyoak, K. J., & Cannon, T. D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition*, *35*, 1445–1455. **DOI:** https://doi.org/10.3758/BF03193614, **PMID:** 18035640

Cho, S., Moody, T. D., Fernandino, L., Mumford, J. A., Poldrack, R. A., Cannon, T. D., et al. (2010). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. *Cerebral Cortex*, *20*, 524–533. **DOI:** https://doi.org/10.1093/cercor/bhp121, **PMID:** 19549622

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., et al. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage*, *14*, 1136–1149. **DOI:** https://doi.org/10.1006/nimg.2001.0922, **PMID:** 11697945

Curtin, J. (2018). lmSupport: Support for linear models (R package version 2.9.13). Retrieved from https://CRAN.R-project.org/package=lmSupport.

Flemming, T. M., Beran, M. J., Thompson, R. K. R., Kleider, H. M., & Washburn, D. A. (2008). What meaning means for same and different: Analogical reasoning in humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and rhesus monkeys (*Macaca mulatta*). *Journal of Comparative*

Psychology, *122*, 176–185. **DOI:** https://doi.org/10.1037/0735-7036.122.2.176, **PMID:** 18489233, **PMCID:** PMC4206216
Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170. **DOI:** https://doi.org/10.1207/s15516709cog0702_3
Gentner, D. (2002). Analogy in scientific discovery: The case of Johannes Kepler. In L. Magnani & N. J. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp. 21–39). New York: Kluwer Academic. **DOI:** https://doi.org/10.1007/978-1-4615-0605-8_2
Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34*, 752–775. **DOI:** https://doi.org/10.1111/j.1551-6709.2010.01114.x, **PMID:** 21564235
Gentner, D. (2016). Language as cognitive tool kit: How language supports relational thought. *American Psychologist*, *71*, 650–657. **DOI:** https://doi.org/10.1037/amp0000082, **PMID:** 27977235
Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306–355. **DOI:** https://doi.org/10.1016/0010-0285(80)90013-4
Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38. **DOI:** https://doi.org/10.1016/0010-0285(83)90002-6
Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature*, *410*, 930–933. **DOI:** https://doi.org/10.1038/35073582, **PMID:** 11309617
Glahn, D. C., Cannon, T. D., Gur, R. E., Ragland, J. D., & Gur, R. C. (2000). Working memory constrains abstraction in schizophrenia. *Biological Psychiatry*, *47*, 34–42. **DOI:** https://doi.org/10.1016/S0006-3223(99)00187-0
Goswami, U. (1992). *Analogical reasoning in children*. Hove, UK: Erlbaum.
Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322. **DOI:** https://doi.org/10.1038/nn1014, **PMID:** 12592404
Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: Semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex*, *20*, 70–76. **DOI:** https://doi.org/10.1093/cercor/bhp081, **PMID:** 19383937
Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2012). Neural correlates of creativity in analogical reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 264–272. **DOI:** https://doi.org/10.1037/a0025764, **PMID:** 22103784
Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development*, *35*, 193–217. **DOI:** https://doi.org/10.1159/000277167
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*, 803–831. **DOI:** https://doi.org/10.1017/S0140525X98001769, **PMID:** 10191879
Hofstadter, D., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. New York: Basic Books.
Holyoak, K. J., & Kroger, J. K. (1995). Forms of reasoning: Insight into prefrontal functions? *Annals of the New York Academy of Sciences*, *769*, 253–264. **DOI:** https://doi.org/10.1111/j.1749-6632.1995.tb38143.x, **PMID:** 8595030
Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press. **DOI:** https://doi.org/10.7551/mitpress/4549.001.0001
Jones, L. L., Kmiecik, M. J., Irwin, J. L., Unsworth, N., & Morrison, R. G. (in preparation). Semantic distance and distracter salience in verbal analogy.
Judd, C. M., McClelland, G. H., & Ryan, C. S. (2017). *Data analysis: A model comparison approach to regression, ANOVA, and beyond* (3rd ed.). New York: Routledge. **DOI:** https://doi.org/10.4324/9781315744131
Kmiecik, M. J., Schauer, G. F., Martinez, D. M., & Krawczyk, D. C. (2016). The similar situations task: An assessment of analogical reasoning in healthy and clinical populations. Poster presented at the Cognitive Neuroscience Society Annual Meeting.
Krawczyk, D. C. (2012). The cognition and neuroscience of relational reasoning. *Brain Research*, *1428*, 13–23. **DOI:** https://doi.org/10.1016/j.brainres.2010.11.080, **PMID:** 21129363
Krawczyk, D. C., Morrison, R. G., Viskontas, I., Holyoak, K. J., Chow, T. W., Mendez, M. F., et al. (2008). Distraction during relational reasoning: The role of prefrontal cortex in interference control. *Neuropsychologia*, *46*, 2020–2032. **DOI:** https://doi.org/10.1016/j.neuropsychologia.2008.02.001, **PMID:** 18355881
Kroger, J. K., Holyoak, K. J., & Hummel, J. E. (2004). Varieties of sameness: The impact of relational complexity on perceptual comparisons. *Cognitive Science*, *28*, 335–358. **DOI:** https://doi.org/10.1207/s15516709cog2803_2
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. **DOI:** https://doi.org/10.1093/cercor/12.5.477, **PMID:** 11950765
Martinez, D. M., Schauer, G. F., Kmiecik, M. J., Kamat, P., Sivess, E., Pires, B., et al. (in preparation). Exploring what constitutes analogical reasoning using a novel test of visual relational similarity.
Miyashita, Y., Higuchi, S.-I., Sakai, K., & Masui, N. (1991). Generation of fractal patterns for probing the visual memory. *Neuroscience Research*, *12*, 307–311. **DOI:** https://doi.org/10.1016/0168-0102(91)90121-E
Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L., et al. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience*, *16*, 260–271. **DOI:** https://doi.org/10.1162/089892904322984553, **PMID:** 15068596
Pederson, T. L. (2019). patchwork: The composer of plots (R package version 1.0.0). Retrieved from https://CRAN.R-project.org/package=patchwork.
Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*, 109–130. **DOI:** https://doi.org/10.1017/S0140525X08003543, **PMID:** 18479531
Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology*, *33*, 43–63. **DOI:** https://doi.org/10.1006/cogp.1997.0659, **PMID:** 9212721
Premack, D. (1983). The codes of man and beasts. *Behavioral and Brain Sciences*, *6*, 125–136. **DOI:** https://doi.org/10.1017/S0140525X00015077
Premack, D. (2010). Why humans are unique: Three theories. *Perspectives on Psychological Science*, *5*, 22–32. **DOI:** https://doi.org/10.1177/1745691609356782, **PMID:** 26162060
Raven, J. C. (1941). Standardization of progressive matrices, 1938. *British Journal of Medical Psychology*, *19*, 137–150. **DOI:** https://doi.org/10.1111/j.2044-8341.1941.tb00316.x
R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.
Revelle, W. (2019). psych: Procedures for psychological, psychometric, and personality research. Retrieved from https://CRAN.R-project.org/package=psych.
Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science*, *316*, 1128–1129. **DOI:** https://doi.org/10.1126/science.1142103, **PMID:** 17525320

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.

Robinson, D., Hayes, A., & Couch, S. (2018). broom: Convert statistical analysis objects into tidy tibbles. Retrieved from https://CRAN.R-project.org/package=broom.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.

Smirnova, A., Zorina, Z., Obozova, T., & Wasserman, E. (2015). Crows spontaneously exhibit analogical reasoning. *Current Biology*, *25*, 256–260. **DOI:** https://doi.org/10.1016/j.cub.2014.11.063, **PMID:** 25532894

Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, *24*, 363–396. **DOI:** https://doi.org/10.1207/s15516709cog2403_2

Urbanski, M., Bréchemier, M.-L., Garcin, B., Bendetowicz, D., de Schotten, M. T., Foulon, C., et al. (2016). Reasoning by analogy requires the left frontal pole: Lesion-deficit mapping and clinical implications. *Brain*, *139*, 1783–1799. **DOI:** https://doi.org/10.1093/brain/aww072, **PMID:** 27076181

Vendetti, M. S., Matlen, B. J., Richland, L. E., & Bunge, S. A. (2015). Analogical reasoning in the classroom: Insights from cognitive science. *Mind, Brain, and Education*, *9*, 100–106. **DOI:** https://doi.org/10.1111/mbe.12080

Volle, E., Gilbert, S. J., Benoit, R. G., & Burgess, P. W. (2010). Specialization of the rostral prefrontal cortex for distinct analogy processes. *Cerebral Cortex*, *20*, 2647–2659. **DOI:** https://doi.org/10.1093/cercor/bhq012, **PMID:** 20156841, **PMCID:** PMC2951846

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. **DOI:** https://doi.org/10.1111/1467-9280.00118

Wechsler, D. (2009). *Wechsler memory scale* (4th ed.). San Antonio, TX: Pearson.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. **DOI:** https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., François, R., Henry, L., & Müller, K. (2018). dplyr: A grammar of data manipulation (R package version 0.7.6). Retrieved from https://CRAN.R-project.org/package=dplyr.