Individual differences in spontaneous analogical transfer

James R. Kubricht[1] · **Hongjing Lu**[1] · **Keith J. Holyoak**[1]

Abstract Research on analogical problem solving has shown that people often fail to spontaneously notice the relevance of a semantically remote source analog when solving a target problem, although they are able to form mappings and derive inferences when given a hint to recall the source. Relatively little work has investigated possible individual differences that predict spontaneous transfer, or how such differences may interact with interventions that facilitate transfer. In this study, fluid intelligence was measured for participants in an analogical problem-solving task, using an abridged version of the Raven's Progressive Matrices (RPM) test. In two experiments, we systematically compared the effect of augmenting verbal descriptions of the source with animations or static diagrams. Solution rates to Duncker's radiation problem were measured across varying source presentation conditions, and participants' understanding of the relevant source material was assessed. The pattern of transfer was best fit by a moderated mediation model: the positive impact of fluid intelligence on spontaneous transfer was mediated by its influence on source comprehension; however, this path was in turn modulated by provision of a supplemental animation via its influence on comprehension of the source. Animated source depictions were most beneficial in facilitating spontaneous transfer for those participants with low scores on the fluid intelligence measure.

✉ James R. Kubricht
kubricht@ucla.edu

[1] Department of Psychology, University of California, 405 Hilgard Ave., Los Angeles, CA 90095, USA

Analogical inference—the application of knowledge about a familiar source system to a novel but structurally similar target system—is critical in scientific discovery (Dunbar & Klahr, 2012) and many other types of creative human activity (Gentner, 2010; Holyoak, 2012; Holyoak & Thagard, 1995). The human capacity for abstract thinking, which is exemplified by analogical reasoning, exceeds that of any other species and plays a significant role in formulating ideas that transcend immediate perception (Penn, Holyoak, & Povinelli, 2008).

It is generally recognized that analogical reasoning involves several subprocesses, most notably retrieval of a related source analog, mapping, inference, and subsequent generalization (e.g., Holyoak, Novick, & Melz, 1994). A number of computational models of these subprocesses have been proposed, including the Structure Mapping Engine (SME; Falkenhainer, Forbus & Gentner, 1989), the Incremental Analogy Machine (IAM; Keane & Brayshaw, 1988; Keane, Ledgeway, & Duff, 1994), the Structured Tensor Analogical Reasoning model (STAR; Halford, Wilson, & Phillips, 1998), and Learning and Inference with Schemas and Analogies (LISA; Hummel & Holyoak, 1997, 2003). A basic empirical finding is that when a source and target are drawn from different knowledge domains and encountered in different contexts, a potentially useful source analog often remains unnoticed. The gap between noticing and actual use of a source analog has been explored most extensively in experiments using Duncker's (1945) radiation problem as the target analog. In this problem, a doctor must find a way to use a radiation ray of varying intensity to destroy an inoperable stomach tumor in a patient. The essence of the problem is that high-intensity rays will destroy healthy tissue when they pass through it on their way to the tumor. Although low-intensity rays do not harm healthy tissue, they are also ineffective in damaging or destroying the tumor. The convergence solution is to apply multiple low-intensity rays to the tumor simultaneously from multiple locations surrounding the target.

Spontaneous analogical transfer

Gick and Holyoak (1980, 1983) found that in the absence of a related source analog, only about 10% of the participants were able to generate the convergence solution to the radiation problem. When a verbal story highly dissimilar to the radiation problem (a story about a general using converging troops to capture a fortress) was presented to participants prior to the target problem, the rate of spontaneously generating convergence solutions increased to about 30%. After receiving an explicit hint to recall the source analog, an approximately 50% additional participants gave the convergence solution, for a total solution rate of roughly 80%. Thus, people often failed to spontaneously notice the relevance of the source in solving the target problem, though they could successfully form mappings and derive inferences when prompted to do so. The difficulty of spontaneously noticing the relevance of distant analogs remains even when more naturalistic materials are employed (Trench & Minervino, 2015).

Subsequent research has established that spontaneous transfer can be facilitated in a number of ways, including choosing a source analog that is relatively similar to the target (Keane, 1987), or one that permits a clear, isomorphic mapping to the target problem (Holyoak & Koh, 1987). In addition, close comparison of multiple source analogs appears to aid in abstracting a more general schema for a class of problems, which in turn fosters later spontaneous transfer (Catrambone & Holyoak, 1989; Gick & Holyoak, 1983; Loewenstein, Thompson, & Gentner, 2003). Direct instructions to search for remote analogs (without specifying any specific domain) can also be effective (Trench, Olguín, & Minervino, 2016). In general, manipulations that encourage attention to shared relations tend to foster transfer of learning (for a review, see Goldwater & Schalk, 2016).

Less work has been done to address the possibility that people systematically differ in their ability to spontaneously notice and use remote analogies. Evidence shows that measures of fluid intelligence, notably Raven's Progressive Matrices (Raven, 1938), predict performance on standardized analogy tests (Snow, Kyllonen, & Marshalek, 1984) as well as analogical mapping in experimental tasks (Vendetti, Wu, & Holyoak, 2014). Fluid intelligence is closely linked to working-memory capacity (see Ackerman, Beier, & Boyle, 2005; Engle, 2002), which has been linked to analogical reasoning in some computational models (e.g., Halford et al., 1998; Hummel & Holyoak, 2003). However, most tests of analogical reasoning present the source and target together (typically in $A{:}B :: C{:}D$ format), so that the need for spontaneous retrieval of the source is eliminated. Relatively few studies have examined individual differences in more complex analogical problem solving (e.g., Antonietti & Gioletta, 1995). There is some evidence that verbal ability predicts spontaneous transfer rates when the source analog is relatively unfamiliar (Corkill & Fager, 1995), and that a measure of math expertise (math SAT score) predicts spontaneous transfer to a math generalization problem after exposure to multiple target analogs (Novick & Holyoak, 1991).

This study is (to our knowledge) the first to investigate whether fluid intelligence scores predict spontaneous analogical transfer. As noted, computational models of analogical reasoning that operate within working-memory constraints imply that the mapping process will be sensitive to working-memory capacity. However, mapping necessarily *follows* the process of accessing a source analog in memory (since by definition mapping must consider the source and target analogs together). It is therefore less clear on theoretical grounds whether working-memory capacity, or the closely related concept of fluid intelligence, will have an impact on *retrieval* of the source analog, and hence on spontaneous transfer.

We further aimed to determine whether the impact of fluid intelligence on spontaneous transfer is mediated by variations in comprehension of the source analog. In addition, we sought to determine whether the influence of fluid intelligence is modulated by an important factor that may facilitate analogical transfer: providing animations to support comprehension of the source analog.

Visuospatial displays and analogical transfer

Research on diagrammatic reasoning has shown that visuospatial representations of solution strategies for mechanical problems can enhance people's ability to infer the principles of operation for physical systems (Hegarty & Stull, 2012), suggesting the importance of display format in acquisition of abstract knowledge. A few studies have shown that static, visual diagrams can be used as source analogs for verbal target problems (Gick, 1985; Gick & Holyoak, 1983). While uninterpreted diagrams generally result in low rates of spontaneous transfer, they can serve as effective analogs following a hint to recall and apply them to a novel target problem (Gick & Holyoak, 1980).

In contrast to the weak effects of static diagrams, there is some evidence that animated displays can facilitate spontaneous transfer (Beveridge & Parkins, 1987; Pedone, Hummel, & Holyoak, 2001). The radiation problem is temporally dynamic, in that the key concepts involve the summation of forces over space and time. The use of physical motion in an animated display may help the learner to focus attention on dynamic relationships (Tversky & Morrison, 2002), which may in turn provide additional retrieval pathways when the target problem is encountered. Day and Goldstone (2011) found that presenting a force-based physical system can prime dynamic mental models, which in turn facilitates spontaneous transfer when solving social problems based on superficially dissimilar dynamic systems.

The potential effectiveness of animation in promoting transfer is consistent with work in fields such as math education, where it has been advocated that students should be encouraged to ground formalisms in concrete situation models—that is, approximate perceptual and dynamic representations of how events described by a problem text occur in the real world (e.g., Koedinger, Anderson, Hadley, & Mark, 1997; Nathan, 1998; Nathan, Kintsch, & Young, 1992; Reed, 2006). Simulations based on situation models have proved effective in promoting better comprehension of relatively simple texts (e.g., Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004). An important finding is that far transfer is often best promoted by using relatively idealized representations (e.g., simplified rather than detailed perceptual displays; for a review, see Fyfe, McNeil, Son, & Goldstone, 2014).

This study aimed to systematically compare the effectiveness of animations and static diagrams (combined with verbal descriptions) in facilitating spontaneous analogical transfer to the radiation problem. In a previous investigation of the impact of animation on analogical transfer, Pedone et al. (2001) presented animations and diagrams without any verbal cover story and did not measure participants' understanding of the source analogs. Evidence shows that the combination of animations and spoken narration is especially effective in increasing understanding of a mechanical system (Mayer, 2009; Mayer & Anderson, 1991). Thus, animations may provide deeper insight into the causal structure of a dynamic system than does a verbal description alone. To assess this possibility, our study included measures of participants' understanding of the source analog. We hypothesized that animations would improve initial understanding of the source and facilitate subsequent spontaneous analogical transfer. The animations we tested were relatively simple, aiming to focus attention on key relations while grounding verbal problem descriptions in a dynamic perceptual representation. By measuring individual differences, we also sought to determine whether the impact of animation on spontaneous transfer differs for people at varying levels of fluid intelligence.

Experiment 1

Method

Participants

A total of 126 (92 female; $N = 42$ per condition) participants were recruited from the Department of Psychology subject pool at the University of California, Los Angeles (UCLA). They were compensated with course credit for their participation and were naïve to the purpose of the experiment. Three participants with less-than-chance performance on the fluid intelligence measure (Raven's score <2) were removed from the analysis.

Source analog stimuli

Animated stimuli were generated using Psychophysics Toolbox Version 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007) in MATLAB, and presented on a CRT monitor at a viewing distance of 70 cm. In each animation, either one or eight cannons fired small or large cannonballs radially inward toward an enemy octagon surrounded by a friendly barrier. The octagon was colored red and had an initial angular width of 0.89°. In the single-cannon scenarios, the octagon's angular width was reduced by 0° and 0.007° after each small- and large-cannonball volley, respectively. In the multiple-cannon scenarios, its width was reduced by 0.040 and 0.056° after each volley. The small and large cannonballs were colored white and had angular widths of 0.10 and 0.20°, respectively. Each cannon was also colored white and had an angular length and width of 1.13 and 0.57°, respectively. The barrier was colored green and had a radial width that subtended 0.25° of visual angle.

Upon impact of each large cannonball, the outer radius of the barrier was reduced by a Gaussian function with maximum magnitude at the point of impact and decreasing magnitude with azimuthal distance. The outer radius was reduced until the barrier was breached, which occurred after 10 large-cannonball volleys. The minimal radial distance to the barrier and each cannon subtended 1.19 and 4.15° of visual angle, respectively. Following each scenario, a message box appeared, indicating whether the cannon(s) succeeded or failed to meet their objective (i.e., whether the enemy octagon was defeated). Elements in each animation were displayed over a black background. Figure 1 illustrates representative diagrams for each scenario, which were chosen from intermediate frames from their respective animations. The number of cannons and the size of the cannonballs in each scenario, along with the level of damage inflicted to the key elements in each display, are indicated in Table 1. The depicted frames were chosen such that the amount of damage inflicted to the barrier and octagon are visually apparent. Videos of the animated scenarios, along with their accompanying spoken monologues, are included in online Supplemental Materials (http://cvl.psych.ucla.edu/moviedemo.html).

Both animations and diagrams were accompanied by spoken monologues recorded on a Blue Yeti USB microphone with a 48 kHz (16 bit) recording sample rate. In each spoken monologue, the number of cannons and the size of the cannonballs were specified in addition to the amount of damage inflicted to the friendly barrier and enemy octagon by the cannon(s). There were four possible levels of damage (none, minor, moderate, and major), which varied across scenarios. The same source analog stimuli were employed in both Experiments 1 and 2 of this study.

Procedure and design

Participants were randomly assigned into one of three conditions (Verbal, Diagram + Verbal, and Animation + Verbal)
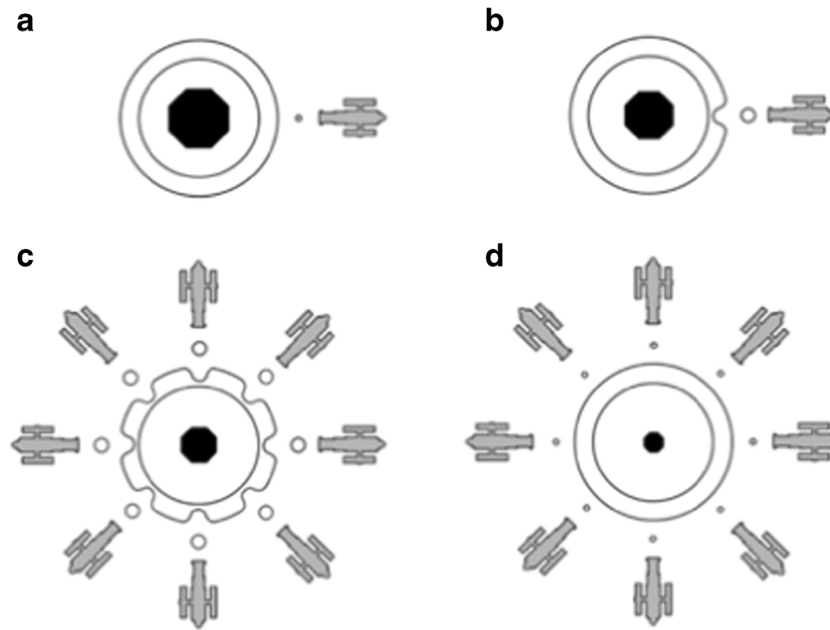
Fig. 1 Illustration of scenario diagrams used in the Diagram + Verbal condition. Participants were presented with the scenarios in sequential order, accompanied by an auditory-verbal explanation of key concepts. **a** In Scenario 1, a single cannon fires small cannonballs and inflicts no damage to either the barrier or octagon. **b** In Scenario 2, a single cannon fires large cannonballs and inflicts minor damage to the octagon and major damage to the barrier. **c** In Scenario 3, multiple cannons fire large cannonballs simultaneously, this time inflicting major damage to both the octagon and barrier. **d** In Scenario 4, multiple cannons fire small cannonballs simultaneously. The converging cannonballs inflict no damage to the barrier and moderate damage to the octagon

reflecting the presentation method for the source analog. In all conditions, an initial instruction page provided pictorial depictions of the general elements in the source analog, which was comprised of four scenarios presented in sequential order. Participants in each condition received an auditory-verbal (spoken monologue) description of each scenario. Those in the Verbal condition were presented with the spoken monologue alone, whereas participants in the Diagram + Verbal and Animation + Verbal conditions received a supplemental diagram or animation, respectively.

The goal of the cannon(s) in each of the four scenarios is to defeat the enemy octagon without inflicting critical damage to any local region of the barrier. The successful scenario is analogous to the convergence solution to Duncker's radiation problem, in which multiple radiation sources are fired at low intensity from multiple locations surrounding the patient's stomach, successfully destroying the tumor without inflicting critical damage to any local region of the surrounding healthy tissue. The relational structure of the cannonball scenario is thus isomorphic to the key relations in the potential solution to the radiation problem.

Questions related to each scenario were presented sequentially following the presentation of all four scenarios. Participants answered two multiple-choice questions per scenario, which assessed their understanding of the level of damage inflicted to the barrier and octagon in each system (e.g., "What level of damage did the cannonballs cause to the friendly barrier in Scenario 1?"). Damage level was directly stated in the spoken monologue for each scenario. After completing the multiple-choice questions, participants were asked to explain why the cannon(s) failed or succeeded in meeting the objective in each scenario (one free-response question per scenario). The question materials are provided in the Appendix. Both multiple-choice and free-response questions were administered using Qualtrics, an online survey environment intended for research and experimental purposes.

Next, each subject completed an abridged, 12-item version of the Raven's Progressive Matrices test (RPM; Arthur, Tubré Paul, & Sanchez-Ku, 1999). In this task, participants viewed a series of 3 × 3 matrices with textured shapes displayed in each cell. The shapes exhibited systematic patterns across the rows and columns of each matrix, and participants were instructed to identify the missing element that completed the pattern without violating the underlying relational rules. The test provides a nonverbal measure of fluid intelligence, allowing us to assess potential individual differences in transfer performance.

Table 1 Number of cannons, size of cannonballs, and level of damage inflicted to key elements in each scenario

Scenario number	Number of cannons	Cannonball size	Barrier damage level	Octagon damage level
1	1	Small	None	None
2	1	Large	Major	Minor
3	8	Large	Major	Major
4	8	Small	None	Moderate

The abridged RPM test also served as a filler task to create a total delay (approximately 15 minutes) between presentation of the source and target analogs. Participants required approximately 11 minutes to complete the RPM test.

Finally, transfer rates to the radiation problem (Duncker, 1945; included in the Supplemental Materials) were measured across varying conditions of source training. Participants were asked to solve Duncker's radiation problem in a two-pass fashion (cf. Gick & Holyoak, 1980). On the first pass, participants received no indication that the previously presented scenarios were related to the target problem. Spontaneous transfer was defined by successful generation of the convergence solution on the first pass without any hint. After participants submitted their answers, the radiation problem was presented again, but this time with an explicit hint prompting participants to recall the cannon scenarios and any solution to the radiation problem they might suggest. Hence, hinted transfer was defined by successful generation of the convergence solution on the second pass, after the explicit hint.

Results

Analogical transfer rates

A set of criteria (adapted from previous research; Gick & Holyoak, 1980) were used to determine whether participants successfully solved the radiation problem either before or after the hint. Solutions were scored according to whether participants conveyed at least two of the three critical ideas underlying the convergence principle (i.e., multiple radiation sources are needed, radiation sources should fire low-intensity rays, and radiation sources should be positioned in different locations surrounding the patient's stomach tumor). Participants were scored as having either (1) solved the radiation problem spontaneously (i.e., without the hint), (2) solved the radiation problem with the hint, or (3) failed to solve the radiation problem. Two undergraduate research assistants naïve to the experimental hypothesis and test conditions independently scored each participant's responses, with an agreement rate of 95% (Cohen's κ = .87). A third researcher broke the tie if the first two researchers disagreed with one another.

Spontaneous transfer rate corresponds to the percentage of participants who produced the convergence solution to the radiation problem before a hint was given to recall the source analog. This measure assesses participants' ability to spontaneously retrieve the source analog and apply their knowledge to a novel target problem following a time delay. Total transfer rate corresponds to the percentage of participants able to solve the radiation problem either before or after a hint was given. Figure 2 depicts percentage of spontaneous transfer before a hint, and total transfer percentage after a hint, for each condition (42 participants per condition).
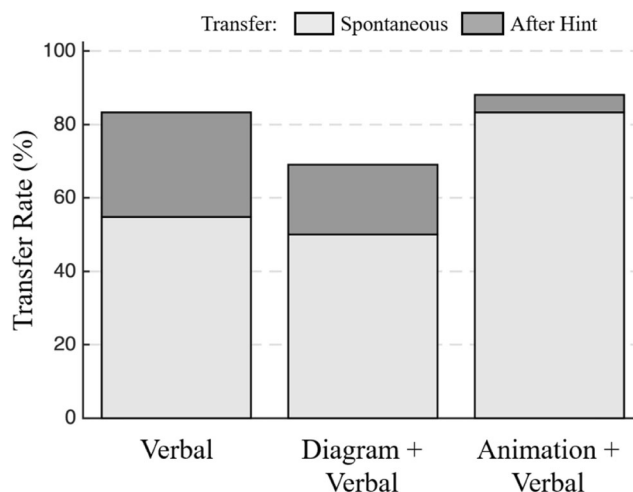

Fig. 2 Spontaneous and total transfer rate (%) for the Verbal, Diagram + Verbal, and Animation + Verbal conditions (Experiment 1)

Spontaneous transfer rates were 55%, 50%, and 83% for the Verbal, Diagram + Verbal, and Animation + Verbal conditions, respectively. The spontaneous transfer rate in the Animation + Verbal condition was reliably greater than that obtained in the Diagram + Verbal, $\chi^2(1) = 10.50$, $\varphi = .35$, $p < .01$, and Verbal conditions, $\chi^2(1) = 8.02$, $\varphi = .31$, $p < .01$, indicating that animated source instruction facilitated spontaneous analogical inference. Spontaneous transfer rate did not differ significantly between the Verbal and Diagram + Verbal conditions, $\chi^2(1) < 1$, $\varphi = .05$, suggesting that the addition of a static pictorial display was not effective in priming the temporally dynamic convergence principle.

Total transfer rates after the hint were 83%, 69%, and 88% for the Verbal, Diagram + Verbal, and Animation + Verbal conditions, respectively. This level of overall performance is consistent with previous findings showing that roughly 80% of college students are able to solve the radiation problem following a hint to think back to a relevant source analog (Gick & Holyoak, 1980, 1983). The total transfer rate in the Animation + Verbal condition exceeded that obtained in the Diagram + Verbal condition, $\chi^2(1) = 4.53$, $\varphi = .23$, $p = .03$, but did not differ from that obtained in the Verbal condition. $\chi^2(1) < 1$, $\varphi = .07$. The total transfer rate in the Diagram + Verbal condition was not statistically different from that in the Verbal condition, $\chi^2(1) = 2.36$, $\varphi = .17$, $p = .12$.

Comprehension of source analog

To assess whether the advantage of the animated source depiction in supporting spontaneous transfer was linked to deeper understanding of the source analog, we evaluated participants' responses to multiple-choice (MC) and free-response (FR) source-understanding questions. The eight MC questions aimed to measure participants' understanding of how the small and large cannonballs interact with the

various elements in the system, and how their forces sum together across scenarios. Each MC question was scored as either correct or incorrect, according to whether the participant selected the correct amount of damage inflicted to the specified element by the cannon(s). The maximum score for MC questions was thus 8 points.

The FR questions assessed participants' understanding of why the cannon(s) failed or succeeded in each scenario. For FR responses, three key principles were chosen for each scenario (see Supplemental Materials). Participants received one point for each correctly conveyed principle and had one point deducted for each incorrect idea they stated. For each question, participants receiving 2 or more points were given a score of 2, participants with one point were given a score of 1, and participants with zero points or less were given a score of 0. The maximum score on the four FR questions was thus 8 points (the same as for MC questions). The FR responses in each scenario were scored by two researchers. If two of the three scorers agreed, their score was used. A third scorer was employed for those responses where the first two disagreed. If the three scorers disagreed, the response was jointly discussed until two researchers agreed on a score. The agreement rate for the first two scorers was 80% (Cohen's κ = .62). The agreement rate for three scorers (i.e., cases where two of the three scorers agreed) was 95%.

FR and MC scores were correlated across participants (r = .47, $p < .001$). For FR scores, participants in the Animation + Verbal condition scored highest, $t(82) = 2.27$, $d = .50$, $p = .03$, Animation versus Diagram, followed by the Diagram + Verbal condition, $t(82) = 2.23$, $d = .49$, $p = .03$, Diagram versus Verbal, and the Verbal condition scored the lowest. For MC scores, participants in the Animation + Verbal condition again scored higher than those in either the Diagram + Verbal, $t(82) = 2.76$, $d = .60$, $p < .01$, or Verbal conditions, $t(82) = 3.07$, $d = .67$, $p < .01$, whereas there was no significant score difference between Diagram + Verbal and Verbal participants, $t(82) < 1$, $d = .11$. The consistent superiority of the animation condition in both FR and MC responses indicates that presence of a supplemental animation promoted deeper understanding of the source analog, relative to other presentation methods (see Fig. 3).

Individual differences in fluid intelligence in relation to spontaneous analogical transfer

To determine whether fluid intelligence influences spontaneous analogical transfer and its interaction with presentation format of the source analog, we performed a median split on participants according to their Raven's score, classifying them as either low (Raven's score <8) or high (Raven's score >8) on the RPM test. Forty-five participants were classified as low RPM (N = 15, 12, and 18 for the Verbal, Diagram + Verbal, and Animation + Verbal conditions, respectively), and 57 participants were classified as high RPM (N = 19 for each condition). The other 24
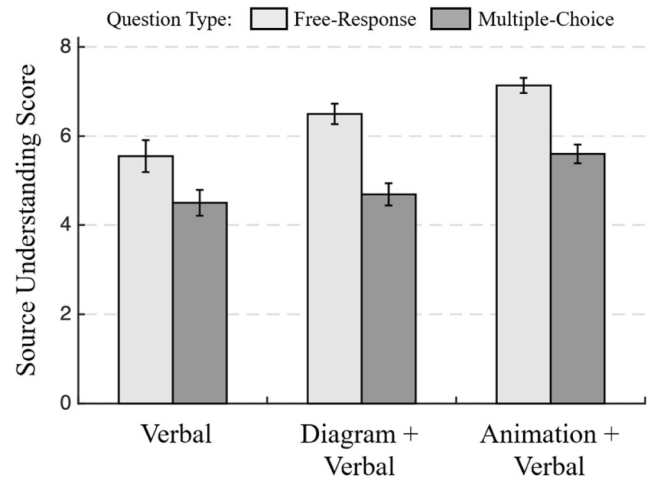

Fig. 3 Free-response (FR) and multiple-choice (MC) source-understanding scores for participants in the Verbal, Diagram + Verbal, and Animation + Verbal conditions (Experiment 1). *Error bars* indicate ± 1 standard error of the mean (*SEM*)

participants, who had median Raven's scores of 8, were dropped from the individual differences analysis as they could not be reasonably classified as either high or low. Data from those participants with median Raven's scores were, however, used in a subsequent moderated mediation analysis, which aggregated data from Experiments 1 and 2 (see Moderated Mediation Analysis section).

We first compared free-response (FR) and multiple-choice (MC) source-understanding scores between the low- and high-RPM groups. Both FR and MC scores in the high-RPM group exceeded those in the low-RPM group, $t(100)$ = 3.66, $d = .73$, $p < .001$ for FR scores; $t(100) = 3.83$, $d = .76$, $p < .001$ for MC scores, indicating superior source understanding for participants scoring relatively high on our fluid intelligence measure. We then compared FR and MC scores between conditions for the low- and high-RPM groups (see Fig. 4). For low-RPM participants, both FR and MC scores in the Animation + Verbal condition exceeded those in the Diagram + Verbal, $t(28) = 3.79$, $d = 1.41$, $p < .001$ for FR scores; $t(28) = 2.09$, $d = .78$, $p < .05$ for MC scores, and Verbal condition, $t(31) = 3.92$, $d = 1.37$, $p < .001$ for FR scores; $t(31)$ = 4.37, $d = 1.53$, $p < .001$ for MC scores, indicating enhanced source understanding for low-RPM participants when the source analog included an animation. For high-RPM participants, MC scores did not differ reliably across the three conditions. However, FR scores in the Animation + Verbal condition exceeded those in the Verbal condition, $t(36) = 2.25$, $d = .73$, $p = .03$, although no difference was observed between FR scores in the Animation + Verbal and Diagram + Verbal conditions, $t(36) < 1$, $d = .12$. FR scores in the Diagram + Verbal condition also exceeded those obtained in the Verbal condition, $t(36) = 2.03$, $d = .66$, $p < .05$, indicating that static diagrams were sufficient to facilitate source understanding (measured by the FR metric) for high-RPM participants.
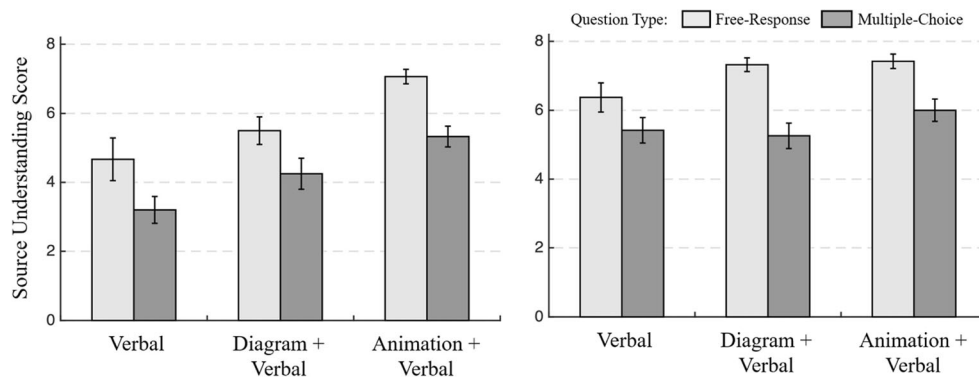
Fig. 4 Free-response (FR) and multiple-choice (MC) source-understanding scores for low-RPM (*left*) and high-RPM (*right*) participants across experimental conditions (Experiment 1). *Error bars* indicate ± 1 *SEM*

Spontaneous transfer rates were then compared across RPM groups, revealing superior performance for those participants scoring high on the fluid intelligence measure, $\chi^2(1) = 4.47$, $\varphi = .21$, $p = .03$. (Because total transfer rates after a hint approached the effective ceiling level, analyses of individual differences were not performed for that measure.) Next, spontaneous transfer rates were compared across conditions for the low- and high-RPM groups (see Fig. 5). For low-RPM participants, the spontaneous transfer rate for the radiation problem in the Animation + Verbal condition exceeded that obtained in the Diagram + Verbal, $\chi^2(1) = 7.75$, $\varphi = .51$, $p < .01$, and Verbal, $\chi^2(1) = 6.64$, $\varphi = .45$, $p = .03$, conditions. No difference was observed in transfer performance between the Diagram + Verbal and Verbal conditions, $\chi^2(1) < 1$, $\varphi = .07$. For high-RPM participants, the spontaneous transfer rate in the Animation + Verbal condition was marginally greater than that in the Diagram + Verbal condition, $\chi^2(1) = 3.64$, $\varphi = .31$, $p = .06$, but did not differ from that in the Verbal condition, $\chi^2(1) = 1.58$, $\varphi = .20$, $p = .21$. The spontaneous transfer rate in the Diagram + Verbal condition also did not differ from that in the Verbal condition, $\chi^2(1) < 1$, $\varphi = .11$.

In summary, animated source depictions were most beneficial in fostering comprehension of the source analog and spontaneous analogical transfer for those participants with low scores on the fluid intelligence measure. After describing

Experiment 2, we will report mediation analyses that further gauge the relationships among animation, source understanding, and spontaneous transfer in analogical discovery.

Experiment 2

The results of Experiment 1 indicate that the advantage of an animated source analog in supporting spontaneous transfer was linked to greater understanding of the source scenarios, measured by our MC and FR questions. One issue that remains open is what, if any, impact the answering of questions about the source may have itself had on spontaneous transfer. The questions provided a measure of source understanding that proved to be predictive of spontaneous transfer. However, to establish a baseline for comparing transfer rates in our studies to those in related studies that did not administer such questions, Experiment 2 explicitly varied whether or not participants answered questions about the source analog.

It is unclear whether answering such questions itself facilitates or impedes transfer, compared to previous studies in which participants' understanding of source analogs was not measured prior to presentation of the radiation problem (e.g., Gick & Holyoak, 1980, 1983; Pedone et al., 2001). On the one hand, responding to questions about the cannon scenarios
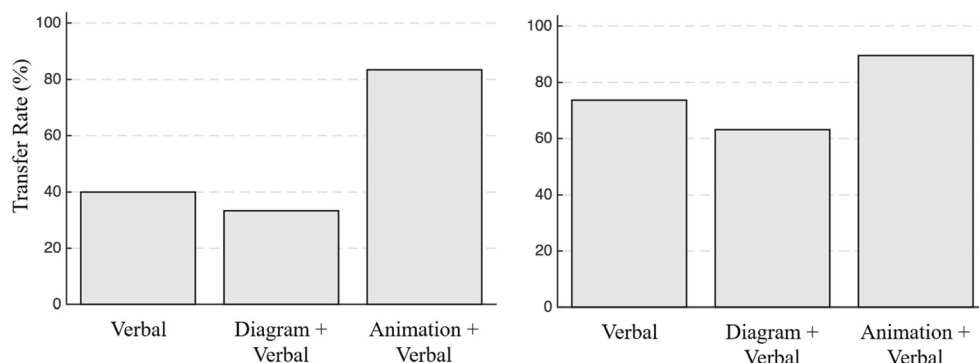

Fig. 5 Spontaneous transfer rate (percentage of convergence solutions generated before hint) for low-RPM (*left*) and high-RPM (*right*) participants across experimental conditions (Experiment 1)

could have helped participants attend to the abstract principle as instantiated in the source analog, thereby strengthening potential retrieval cues for subsequent recall and transfer to the radiation problem. On the other hand, the questions probed domain-specific details about the source analog (unlike questions that elicit comparisons of multiple analogs to one another; cf. Gick & Holyoak, 1983). Although performance on the scenario-understanding questions predicted abstract understanding of the underlying principle, it is possible that answering these questions may have strengthened memory for scenario-specific details at a featural level that would not be useful for solving the radiation problem. Thus, the aim of Experiment 2 was to determine what effect, if any, administering source-understanding questions has on analogical transfer to the radiation problem. In addition, we wanted to increase the overall power of the basic design to obtain enough data to perform a mediation analysis.

Method

Participants

A total of 240 participants (159 female; $N = 60$ per condition) were recruited from the Department of Psychology subject pool at UCLA. They were compensated with course credit for their participation and were naïve to the purpose of the experiment. One participant with less-than-chance performance on the fluid intelligence measure (Raven's score <2) was removed from the analysis.

Procedure and design

The Diagram + Verbal condition was dropped in Experiment 2, leaving two levels for the between-subjects factor of source presentation format (Verbal and Animation + Verbal). Participants were randomly assigned into two conditions of a second between-subjects factor (Questions and No Questions), reflecting whether or not source-understanding questions were asked following presentation of the source scenarios. We recognized that eliminating the questions would decrease the time delay between presentation of the source scenarios and the radiation problem. Since participants in Experiment 1 spent approximately 218 seconds answering source-understanding questions, a delay of 18 seconds was introduced following each RPM item for participants in the No Questions condition (yielding a total delay of approximately 15 minutes). The time required for participants to complete the experiment did not differ between the Questions and No-Questions groups, $t(238) = .48$, $d = .06$, $p = .64$, nor did participants' Raven's scores, t(238) = 1.27, $d = .16$, $p = .21$. In all other respects the procedure and design were identical to those of Experiment 1.

Results

Analogical transfer rates

The criteria for scoring participants' responses to the radiation problem were the same as in Experiment 1. The agreement rate between the two researchers was 88% (Cohen's κ = .77), and a third researcher again broke the tie if the first two researchers' scores disagreed with one another. Because the interaction between source presentation format and presence of questions was not significant in a binary logistic regression model predicting spontaneous and total transfer, we proceeded to interpret the main effects of source presentation format and asking source-understanding questions.

Replicating the major finding of Experiment 1, both spontaneous and total transfer rates in the Animation + Verbal condition exceeded those in the Verbal condition, 83% versus 58%, $\chi^2(1) = 18.15$, $\varphi = .28$, $p < .001$, for spontaneous transfer; 90% versus 73%, $\chi^2(1) = 12.06$, $\varphi = .22$, $p < .001$, for total transfer. Spontaneous transfer rates were 63% in the Questions condition and 79% in the No-Questions condition, indicating that the presence of source-understanding questions appeared to impede spontaneous transfer performance, $\chi^2(1) = 8.07$, $\varphi = .18$, $p < .01$. Total transfer rates did not differ significantly between the No-Questions and Questions conditions, 85% versus 78%, $\chi^2(1) = 2.22$, $\varphi = .10$, $p = .14$.

These findings suggest that answering questions about the source scenarios may strengthen memory for source-specific information that is not helpful for spontaneously transferring abstract knowledge between the source analog and the radiation problem. Such source-specific information would increase the semantic distance between the source and target, increasing the difficulty of subsequent transfer. Consequently, the net effect of asking questions was to impair spontaneous retrieval of the source.

Moderated mediation analysis

To assess the impact of individual differences in fluid intelligence on spontaneous transfer, directly and via an influence on source comprehension—and how animation modulates these relationships—we conducted mediation and moderation analyses on combined data from Experiment 1 and the Questions condition of Experiment 2. Hence, the analysis was based on observations from a total of 204 participants. These analyses treated fluid intelligence, as measured by RPM score, as a continuous variable, thereby avoiding the pitfalls associated with creating a binary dichotomous variable (i.e., classifying participants as either low or high RPM). A single source-understanding score was calculated for each participant by summing together FR and MC scores, and a dummy coded variable was created to indicate presence of animation (1 if animations were present, 0 if absent). Both RPM and

summed source-understanding scores were standardized prior to the mediation analysis.

As shown in Fig. 6a, the relationship between fluid intelligence and spontaneous analogical transfer was mediated by source understanding. Moreover, the indirect effect of fluid intelligence on spontaneous transfer through source understanding was moderated by provision of a supplemental animation with the source analog. We compared this moderated mediation model to a baseline mediation model including an indirect effect of presence of animation on spontaneous transfer through source understanding, but omitting the influence of individual differences in fluid intelligence (see Fig. 6b). Log likelihoods were calculated using predicted source-understanding scores, and a likelihood ratio test was conducted to compare the goodness of fit for each model. Results revealed a superior goodness of fit for the moderated mediation model that included the individual difference variable, $\chi^2(2) = 14.84$, Cramer's $V(1) = .35$, $p < .001$.

We tested the significance of links in the baseline and moderated mediation models using bootstrapping procedures. Standardized indirect effects were computed for each of 50,000 bootstrapped samples, and the 95% confidence interval was computed by determining the indirect effects at the 2.5th and 97.5th percentiles. If the bootstrapped confidence interval does not include zero, the corresponding effect is significant at $p = .05$. For the baseline mediation model, the bootstrapped standardized indirect effect was statistically significant ($ab = .80$, $CI_{95} = [.45, 1.28]$), as was the relationship between source understanding and spontaneous transfer ($b = 1.00$, $CI_{95} = [.59, 1.40]$). The relationship between presence of animation and source understanding was also statistically significant ($a = .81$ $CI_{95} = [.55, 1.06]$), whereas the direct effect of presence of animation on spontaneous transfer was not ($c' = .54$, $CI_{95} = [-.15, 1.23]$).

For the moderated mediation model, the bootstrapped standardized indirect effects in the presence and absence of animation were statistically significant ($ab = .20$, $CI_{95} = [.06, .39]$ and $ab = .61$, $CI_{95} = [.31, 1.00]$, respectively). The relationship between source understanding and spontaneous transfer rate was also significant ($b = 1.01$, $CI_{95} = [.61, 1.41]$). However, the direct effect of Raven's score on spontaneous transfer was not reliable ($c' = .29$, $CI_{95} = [-.06, .64]$). To test whether presence of animation significantly moderated this mediation effect, we calculated the 95% confidence interval of the standardized index of moderated mediation (i.e., the difference between the indirect effects in the presence and absence of
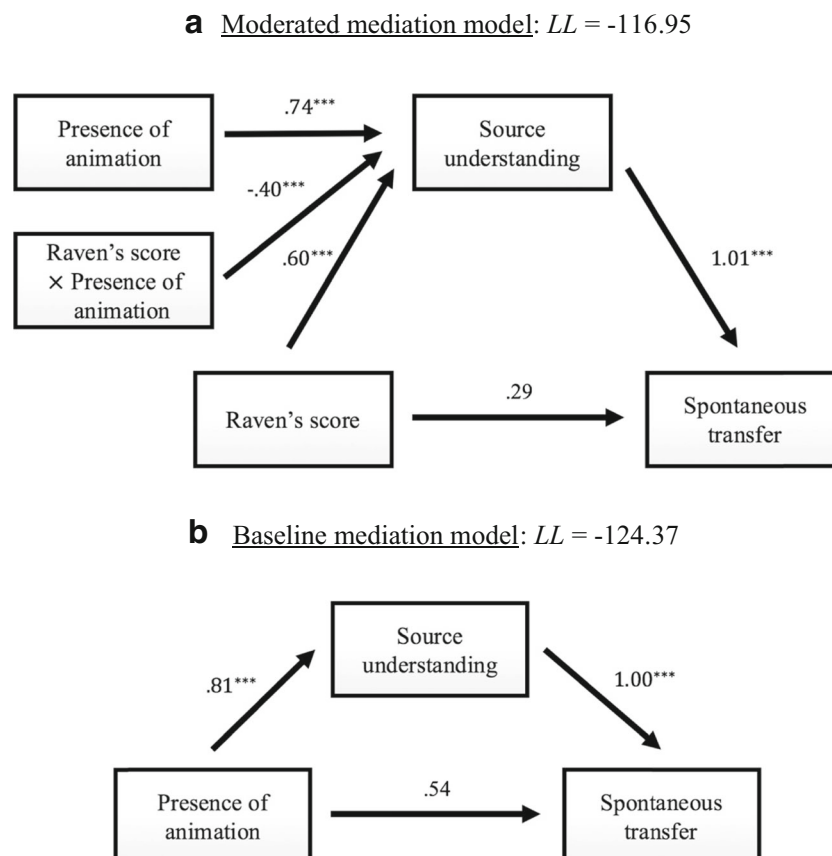
a   Moderated mediation model: $LL = -116.95$

b   Baseline mediation model: $LL = -124.37$

Fig. 6 Standardized regression coefficients for **a** the indirect effect of Raven's score on spontaneous transfer through source understanding moderated by presence of animation and **b** the indirect effect of presence of animation on spontaneous transfer through source understanding. The log likelihood (LL) value indicates the goodness of fit for each model; values closer to zero correspond to higher likelihoods

animation). We found that the standardized index of moderated mediation was statistically significant ($\Delta ab = .40$, $CI_{95} = [.12, .79]$), indicating that the indirect effect of fluid intelligence on spontaneous transfer through source understanding was moderated by the presence of animation.

In Experiment 1, we found that supplemental animations were most beneficial for those participants scoring low on the fluid intelligence measure, suggesting a stronger relationship between Raven's score and source understanding in the absence of animation. Results from the moderated mediation analysis indicate that both presence of animation and Raven's score had a significantly positive impact on source understanding ($a_{Animation} = .74$, $CI_{95} = [.52, .97]$ and $a_{Ravens} = .60$, $CI_{95} = [.43, .77]$). However, the impact of Raven's score on source understanding was attenuated by the presence of animation, as demonstrated by the significant interaction term between the two variables ($a_{Animation*Ravens} = -.40$, $CI_{95} = [-.63, -.18]$). Since source understanding predicts spontaneous transfer, we would expect the indirect effect of Raven's score through source understanding to be greatest in the absence of animation, which is supported by the significant index of moderated mediation. These results provide converging evidence that provision of supplemental animations facilitates spontaneous transfer by increasing comprehension of source materials, and that this benefit is greatest for participants scoring relatively low on the measure of fluid intelligence.

General discussion

In two experiments, we found that animated source analogs yielded greater spontaneous transfer than either diagrammatic or purely verbal source analogs. These findings are consistent with those of Beveridge and Parkins (1987) and Pedone et al. (2001). This study went beyond prior work by measuring the impact of both individual differences in fluid intelligence, and varying presentation conditions, on source understanding. We found that providing a supplemental animation of the source analog led to superior source understanding as well as greater spontaneous analogical transfer. In addition, the presence of animation was especially beneficial for participants who scored lower on a measure of fluid intelligence. To the best of our knowledge, this is the first demonstration that fluid intelligence is a predictor of spontaneous analogical transfer (as opposed to prompted analogical mapping; see Vendetti et al., 2014).

Animation may aid transfer in at least two ways. First, animated displays serve to ground a verbal problem in a dynamic perceptual experience, creating the benefits associated with multimedia instruction (e.g., Reed, 2006). Second, if the animated display is relatively sparse with respect to visual details, then it may focus attention selectively on the relational aspects of the source analog (see Goldwater & Schalk, 2016). In the latter respect, the benefit of animation in promoting an abstract understanding of the source may be similar to the benefit of comparing multiple source analogs, which appears to foster induction of a more abstract schema (Gick & Holyoak, 1983; Novick & Holyoak, 1991). Viewing supplemental animated displays—like comparison of multiple source analogs—may foster richer relational encodings (Pedone et al., 2001). The dynamic nature of animation may be especially effective in inducing a schema for problems involving the application of forces over time and space. It is possible that an animated source induces internal scan patterns across a mental image of the target problem, which in turn facilitates transfer (analogous to benefits conveyed by certain patterns of over eye movements across a diagram of the radiation problem; Grant & Spivey, 2003; Thomas & Lleras, 2007).

Of course, we have only explored the impact of one animated display. Future work should systematically vary the visual richness of such displays to assess whether sparse displays are especially effective in promoting transfer (cf. Fyfe et al., 2014). It is also of interest that the overall rate of spontaneous transfer was substantially higher for the source used in the present study (even in the Verbal-only condition) than in studies using the military source analog (Gick & Holyoak, 1980, 1993). Although there are many differences between the two source analogs, one that may be especially important is that relative to the military story, the "cannonball" scenario used here is less rich in details irrelevant to the convergence solution. Future research could fruitfully examine the general role of degree of specific detail as a factor that influences analogical transfer.

A moderated mediation analysis revealed that fluid intelligence influences spontaneous transfer indirectly, by increasing source comprehension. A possible mechanism is that participants with higher Raven's scores tend to focus more attention on the underlying relations in the source scenario, essentially adopting a "relational set" (Vendetti et al., 2014). The result is a deeper understanding of the source, based on relational concepts that provide potential retrieval paths once the target problem is encountered. This explanation of how fluid intelligence may facilitate spontaneous analogical transfer is consistent with recent evidence that fluid intelligence as measured by the Raven's Matrices test yields benefits that are largely mediated by the use of effective cognitive strategies (Gonthier & Thomassin, 2015; see also Dunlosky & Kane, 2007).

Animation reduces the impact of fluid intelligence on source comprehension, as it enables those with lower fluid intelligence scores to "catch up" in their source understanding. Our findings suggest that dynamic displays may be especially useful in teaching relational concepts to students at lower ability levels. Providing these students with more robust representations of a novel source analog may reduce working-memory demands and encourage a relational set, thereby fostering analogical transfer.

Future theoretical work should focus on how dynamic schemas can be learned from animations, perhaps by extending current theories of relation learning (e.g., Lu, Chen & Holyoak, 2012) to analogical transfer. In addition, across a

broad range of topics, the potential value of animations as a teaching tool that can be effectively coupled with analogical examples deserves to be explored further.

Appendix

Source-understanding questions (multiple choice)

1. What level of damage did the cannon balls cause to the friendly barrier in Scenario #[1–4]
2. What level of damage did the cannon ball(s) cause to the enemy octagon in Scenario #[1–4]

Source-understanding questions (free response)

1. Why was the cannon [were the cannons] unable to defeat the enemy octagon in Scenario #[1–3]?
2. Why were the cannons able to defeat the enemy octagon in Scenario #[4]?

Radiation problem (story)

Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient, but unless the tumor is destroyed, the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays reach the tumor all at once at a sufficiently high intensity, the tumor will be destroyed. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities, the rays are harmless to healthy tissue, but they will not affect the tumor either. What type of procedure might be used to destroy the tumor with the rays, and at the same time avoid destroying the healthy tissue?

Radiation problem (question before hint)

Please write down as many possible solutions you can think of, without worrying about whether solutions would really work. No medical expertise is required.

Radiation problem (question after hint)

Think back to the scenarios you observed earlier and write down any solution to the radiation problem they may suggest. It's okay to repeat a solution you gave earlier.

Scoring criteria (key principles)

Scenario 1:

–   The small cannonballs inflicted no damage to the octagon.
–   There was only one cannon firing cannonballs at the enemy octagon as opposed to multiple cannons.
–   The small cannonballs inflicted no damage to the friendly barrier.

Scenario 2:

–   The large cannonballs inflicted minor/some damage to the octagon.
–   There was only one cannon firing cannonballs at the enemy octagon as opposed to multiple cannons.
–   The large cannonballs inflicted major/substantial damage to the friendly barrier.

Scenario 3:

–   The multiple large cannonballs inflicted substantial/major damage to the octagon (by *converging* onto their target *simultaneously*).
–   There were multiple cannons firing cannonballs at the enemy octagon as opposed to a single cannon.
–   The large cannonballs inflicted substantial/major damage to the friendly barrier.

Scenario 4:

–   The multiple small cannonballs inflicted some/moderate damage to the octagon (by *converging* onto their target *simultaneously*).
–   There were multiple cannons firing cannonballs at the enemy octagon as opposed to a single cannon.
–   The small cannonballs inflicted no damage to the friendly barrier.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131,* 30–60.

Antonietti, A., & Gioletta, M. A. (1995). Individual differences in analogical problem solving. *Personality and Individual Differences, 18,* 611–619.

Arthur, W., Tubré, T., Paul, D. S., & Sanchez-Ku, M. L. (1999). College-sample psychometric and normative data on a short form of the Raven's Advanced Progressive Matrices test. *Journal of Psychoeducational Assessment, 17,* 354–361.

Beveridge, M., & Parkins, E. (1987). Visual representation in analogical problem solving. *Memory and Cognition, 15*(3), 230–237.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1147–1156.

Day, S. B., & Goldstone, R. L. (2011). Analogical transfer from a simulated physical system. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 551–567.

Dunbar, K. N., & Klahr, D. (2012). Scientific thinking and reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford handbook of thinking and reasoning* (pp. 701–718). New York: Oxford University Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58* (Whole No. 270).

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategies to working memory span: A comparison of strategy assessment methods. *Quarterly Journal of Experimental Psychology, 60,* 1227–1245.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19–23.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41,* 1–63.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instructions: A systematic review. *Educational Psychology Review, 26,* 9–25.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775.

Gick, M. L. (1985). The effect of a diagram retrieval cue on spontaneous analogical transfer. *Canadian Journal of Psychology, 39*(3), 460–466.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*(3), 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1–38.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology, 96,* 424–436.

Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin, 142,* 729–757.

Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices. *Journal of Experimental Psychology: General, 144,* 916–924.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14,* 462–466.

Halford, G. S., Wilson, W. H., & Phillips, W. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental and cognitive psychology. *Behavioral Brain Sciences, 21,* 803–831.

Hegarty, M., & Stull, A. T. (2012). Visuospatial thinking. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford handbook of thinking and reasoning* (pp. 606–630). New York: Oxford University Press.

Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford handbook of thinking and reasoning* (pp. 234–259). New York: Oxford University Press.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15*(4), 332–340.

Holyoak, K. J., Novick, L. R., & Melz, E. R. (1994). Component processes in analogical transfer: Mapping, pattern completion, and adaptation. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory, vol. 2: Analogical connections.* Norwood: Ablex.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought.* Cambridge: MIT Press.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104,* 427–466.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110,* 220–264.

Keane, M. T. (1987). On retrieving analogues when solving problems. *Quarterly Journal of Experimental Psychology, 39A,* 29–41.

Keane, M. T., & Brayshaw, M. (1988). The incremental analogy machine: A computational model of analogy. In D. Sleeman (Ed.), *Third European working session on learning* (pp. 53–62). London, UK: Pitman.

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science, 18,* 287–334.

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?* Tutorial session presented at the 30th European Conference on Visual Perception, Arezzo, Italy.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8,* 30–43.

Loewenstein, J., Thompson, L., & Gentner, D. (2003). Analogical learning in negotiation teams: Comparing cases promotes learning and transfer. *Academy of Management Learning and Education, 2,* 119–127.

Mayer, R. E. (2009). *Multimedia learning.* Cambridge: Cambridge University Press.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83,* 484–490.

Nathan, M. J. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments, 5,* 135–160.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra word problem comprehension and its implications for the design of computer learning environments. *Cognition and Instruction, 9*(4), 329–389.

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 510–520.

Pedone, R., Hummel, J. E., & Holyoak, K. J. (2001). The use of diagrams in analogical problem solving. *Memory and Cognition, 29*(2), 213–221.

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.

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

Reed, S. K. (2006). Cognitive architectures for multimedia learning. *Educational Psychologist, 41,* 87–98.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence, vol. 2.* Hillsdale: Erlbaum.

Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin and Review, 14,* 663–668.

Trench, M., & Minervino, R. (2015). The role of surface similarity in analogical retrieval: Bridging the gap between the naturalistic and experimental traditions. *Cognitive Science, 39,* 1292–1319.

Trench, M., Olguín, V., & Minervino, R. (2016). Seek, and ye shall find: Differences between spontaneous and voluntary analogical retrieval. *Quarterly Journal of Experimental Psychology, 69,* 698–712.

Tversky, B., & Morrison, J. B. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57,* 247–262.

Vendetti, M., Wu, A., & Holyoak, K. J. (2014). Far out thinking: Generating solutions to distant analogies promotes relational thinking. *Psychological Science, 25,* 928–933.