

Promoting Spontaneous Analogical Transfer: The Role of Category Status

Sean Snoddy (ssnoddy1@binghamton.edu)

Kenneth J. Kurtz (kkurtz@binghamton.edu)

Department of Psychology, Binghamton University (SUNY)
Binghamton, NY 13902 USA

Abstract

Analogical comparison promotes spontaneous transfer by encouraging a more abstract representation that may be easier to retrieve. The category status hypothesis states that: if knowledge is represented as a relational category, it is easier to activate as a result of categorizing (as opposed to cue-based reminding). To investigate these two pathways to analogical transfer, participants were assigned to different study conditions: 1) standard comparison of two analogs; 2) standard comparison followed by a second comparison of two new analogs; or 3) a guided category-building task based on sequential summarization. Category-building showed a reliably higher rate of spontaneous transfer during an analogical problem solving task than standard comparison (numerically higher than double-comparison). Another experiment measured spontaneous reminders to cues on the basis of matching structure. Category-building showed a reliable advantage over both comparison conditions. This supports categorization as a novel pathway to spontaneous transfer by enhancing retrieval of structurally similar information.

Keywords: concepts and categories; analogy; problem solving; comparison; transfer

General Introduction

People are able to transfer prior knowledge to solve problems in a superficially dissimilar context. Gick and Holyoak (1980) demonstrated that individuals who encoded a base passage, which described how a general captured a fortress by dividing an army into small groups of soldiers that simultaneously attacked the fortress from various angles, were able to transfer the passage's solution to solve an isomorphic target problem about how a doctor could destroy a tumor with a ray of radiation. The prevailing cognitive account (henceforth abstraction account) explains knowledge transfer across domains (e.g., military strategy to medical treatment) in terms of analogy. According to the abstraction account, transfer involves a mapping process where distinct superficial information between analogs is filtered out (e.g., general and doctor), and similar relationships between analogs are placed into correspondence (e.g., simultaneous application) (Gentner, 1983; Gick & Holyoak, 1980). This mapping process allows for candidate inferences from the structure of the base to fill in missing predicates of the target problem, which allows for a solution to be devised (e.g., lower intensity rays simultaneously converging on a tumor from multiple locations) (Gentner, 1983; Gick & Holyoak, 1980).

Despite the capacity for analogical transfer, individuals often fail to spontaneously transfer knowledge from a single

base analog to solve a problem in a different domain without an explicit hint about the base's relevance (Gick & Holyoak, 1980). This failure is known as the paradox of similarity-based retrieval: superficially similar information to the target problem (e.g., other medical problems) is favored during retrieval, even though inferences require structural overlap between the base and target (e.g., the concept of convergence) (Gentner, Ratterman & Forbus, 1993; Holyoak & Koh, 1987; Ross, 1987). With this paradox in mind, the key to understanding how spontaneous transfer occurs is to determine what promotes structure-based retrieval.

Some key findings of the abstraction account are that comparison is an effective way to learn (Alfieri, Nokes-Malach & Schunn, 2013) and to promote spontaneous transfer (Gentner, Loewenstein & Thompson, 2003; Gick & Holyoak, 1983). During comparison, cases with matching structure are presented side-by-side and participants are prompted to consider the similarities between them. This facilitates a mapping process to occur during encoding that is similar to the one that occurs during transfer, which highlights commonalities between cases and promotes the formation of an abstract schema via filtering out surface-level mismatches (Markman & Gentner, 2000). These abstract schemas are more accessible in memory than representations of specific cases due to a lack of superficial mismatches with targets (Forbus, Gentner & Law, 1995; Gick & Holyoak, 1983). The heightened accessibility of abstract schemas facilitates retrieval of structurally similar matches during the memory search triggered by an opportunity to spontaneously transfer knowledge.

While the abstraction account focuses on the role of schema abstraction in spontaneous transfer, the type of materials used in these studies can also be viewed as embodying relational categories (Gentner & Kurtz, 2005). Categorization may provide another pathway to structure-based retrieval. The category status hypothesis predicts that when knowledge is represented in the form of a category, it is fluidly accessed and applied (Kurtz & Honke, submitted). There are three important aspects that contribute to the development of category status. First, *category intension* is knowledge of the category defining structure. This may be similar to an abstract schema, and could be conferred through the comparison process (Goldwater & Schalk, 2016). Second, *category extension* is knowledge of specific members and non-members of a category as well as how to differentiate between them. Third, categorization confers experience in *bi-directional mapping between generic knowledge of the category and specific cases*. Thus, the key

claim of the category status hypothesis is as follows: when knowledge is represented as a psychological category, construal of a target stimulus as a category member facilitates direct activation of the category-defining concept in semantic memory. The key difference is that the category status hypothesis predicts direct activation of structurally relevant matches as opposed to a filtering out of superficial mismatches followed by an evaluation for any structural similarity.

Initial support for the category status hypothesis comes from Kurtz and Honke (submitted) who had participants learn a relational principle either through category construction or a single comparison opportunity. In the category construction task, participants formed two categories out of three examples of a principle and three alignably different examples. This was compared to the standard version of the comparison task that received two cases presented side-by-side and provided a similarity rating as well as an explanation of the similarities between cases. Category construction led to a higher rate of spontaneous transfer than the comparison task, which suggested that promoting category status may provide a novel pathway to spontaneous transfer (Kurtz & Honke, submitted).

The goals of the present work are: 1) to further evaluate the plausibility of the category status hypothesis as an alternative account of spontaneous transfer, and 2) to further explore the abstraction account. To address the first goal, a novel category building task based on the sequential summarization of cases was used to promote category status. This task is often used as a control to comparison, and does not confer the same level of abstraction-based transfer benefits (Catrambone & Holyoak, 1989; Gentner et al., 2003; Rittle-Johnson & Star, 2007). If sequential summarization can be combined with additional supports that promote category status, then it should become an effective way to promote transfer. The categorization supports that were integrated with the core summarization task were: 1) summarization of category-membership-relevant aspects of multiple cases, 2) identification of each case with a shared category label, and 3) a description of the category after encountering all cases. If these supports contribute to the development of category status, a categorization-based summarization task should become an effective way to promote transfer.

The second goal of the present work sought to further understand the abstraction account. Prior analogical comparison research has largely focused on the effects of a single comparison opportunity on transfer success. The effect of an additional comparison opportunity was explored, which should improve schema abstraction by providing additional surface-level mismatches to filter out. The additional comparison opportunity also serves as a control for case exposure in the category-building condition.

Experiment 1

Experiment 1 used the analogical transfer paradigm (Gick & Holyoak, 1980, 1983) to assess the impact of categorization-

based summarization (category-building), the standard version of the comparison task (single comparison) (cf. Catrambone & Holyoak, 1989; Gentner et al., 2003; Gick & Holyoak, 1983), a standard comparison task that is repeated a second time with novel cases (double comparison), and a baseline condition on spontaneous transfer performance. Spontaneous transfer success is contingent upon being able to both spontaneously access and retrieve relevant knowledge from memory as well as apply that knowledge to devise a solution (Gick & Holyoak, 1983). The use of both spontaneous and hint-aided transfer assessments allows for the differentiation of the relative impact that each study condition has on application ability and retrieval.

Both the abstraction account and category status hypothesis make different predictions about what type of task will improve the retrieval process that underlies successful spontaneous transfer. The predictions for spontaneous transfer are as follows: 1) all study conditions will promote transfer (i.e., result in a higher rate of transfer than baseline), 2) the category-building condition will result in a higher rate of transfer than both comparison conditions, and 3) double comparison will result in a higher rate of transfer than single comparison. Neither account makes explicit predictions about application ability, so hint-aided transfer performance is exploratory.

Method

Participants A total of 355 undergraduate students from Binghamton University participated for course credit. Hint-aided transfer data from seven participants were excluded due to a failure to complete the assessment in the allotted time.

Materials and Design The materials consisted of both study cases and the transfer problem. All materials demonstrated the principle “problem-as-a-solution”: *when a large-scale event causes a large amount of damage, the event can be mitigated by repeatedly causing it on a small scale and incurring minor damage each time*. The study cases were all from the domain of natural disasters. The passages were a single paragraph that consisted of a description of the problem followed by the solution.

The transfer problem involved the prevention of cybercrime. It contained a similar description of the problem as the study cases. However, the solution was replaced with an open-ended question about how the threat of computer hackers could be minimized. The general formatting of the transfer problem was different from the study tasks to make the separate phases appear unrelated. The transfer problem was presented twice. It was first presented under the guise of a new experiment about problem solving (spontaneous transfer). The same problem was presented again with a hint for participants to use their knowledge from the study phase (hint-aided transfer).

Procedure Prior to the experiment, all participants were informed that they would take part in multiple experiments,

then were randomly assigned to one of the study tasks. Participants in the single comparison condition received two study cases, which were referred to as solved problems, presented side-by-side. Participants were informed that there were important ways that the solved problems were alike and were asked to consider the similarities between them. After reading the cases, participants provided a similarity rating on a five-point scale that ranged from ‘not at all similar’ to ‘very similar’. Participants were then asked to describe the similarities and differences they considered when making their similarity judgement. The double-comparison task was similar to single comparison, with the only difference being a second comparison opportunity for two novel cases. During this second comparison opportunity, participants were asked to consider the important ways in which all four of the solved problems were alike, provide a similarity rating for the two additional cases, and describe the similarities used to create the rating.

In the category-building condition, participants were given instructions that explicitly stated the task involved learning about a category, and then were sequentially presented with all four cases. Upon presentation of each case, participants were provided with the label to identify the case as a member of the category (e.g., “Here is an example of the Tongo category”) and were asked to summarize the relevant information to category membership. After completion of the final summarization task, participants were asked to provide a description of the category they had just learned.

Following the study task, participants were told that they were beginning a new experiment on problem solving, and proceeded to the spontaneous transfer assessment. Since the baseline condition was meant to establish chance production of transfer solutions to the problem, participants in that condition only received this assessment. After completion of the spontaneous transfer assessment, participants were then given another chance to solve the problem with an explicit hint to use their knowledge from the study phase. Both of these assessments presented participants with the transfer problem, and asked them to devise a solution. For each transfer opportunity, participants were allowed to provide multiple solutions to the problem. Each of the proposed solutions was scored by the first author blind to condition. If at least one of a participant’s proposed solutions demonstrated the target principle, that participant was coded as a transfer success.

Results and Discussion

Hint-aided Transfer Hint-aided transfer performance reflects participants’ ability to retain and apply the knowledge from the learning tasks. To evaluate hint-aided transfer performance, a logistic regression model (R Core Team, 2016) was built to predict hint-aided success with condition. There were no significant differences between category-building and single comparison ($\beta = 0.274$, $SE = 0.313$, Wald $Z = 0.872$, $p = .383$) or double-comparison ($\beta = -0.174$, $SE = 0.318$, Wald $Z = -0.547$, $p = .585$).

Additionally, there were no significant differences between double-comparison and single-comparison ($\beta = 0.447$, $SE = 0.313$, Wald $Z = 1.426$, $p = 0.154$) (see Table 1).

Table 1: Hint-aided transfer performance.

Learning Task	% Transfer (N)	95% C.I.	N
Single-Comparison	55% (47)	45% - 66%	85
Double-Comparison	66% (58)	55% - 75%	88
Category-Building	62% (52)	51% - 72%	84

Spontaneous Transfer Spontaneous transfer success was modelled using a logistic regression with condition as the predictor. Both the category-building ($\beta = 1.68$, $SE = 0.433$, Wald $Z = 3.88$, $p < .001$) and double-comparison ($\beta = 1.1$, $SE = 0.449$, Wald $Z = 2.453$, $p = .014$) conditions resulted in a higher rate of transfer than baseline. However, the single comparison condition was not significantly different from baseline transfer performance ($\beta = 0.407$, $SE = 0.491$, Wald $Z = 0.828$, $p = .408$) (see Table 2).

The category-building condition led to a significantly higher rate of spontaneous transfer than single comparison ($\beta = 1.274$, $SE = 0.393$, Wald $Z = 3.239$, $p < .01$). However, there was no significant difference between category-building and double comparison ($\beta = 0.579$, $SE = 0.339$, Wald $Z = 0.171$, $p = .088$). There was no significant difference between double and single comparison ($\beta = 0.695$, $SE = 0.411$, Wald $Z = 1.692$, $p = .091$) (see Table 2). To account for slight numeric differences in hint-aided transfer, a more conservative analysis was done that included only participants with hint-aided transfer success to clearly reflect differences in the retrieval process. The same pattern of results was observed.

Table 2: Spontaneous transfer performance.

Learning Task	% Transfer (N)	95% C.I.	N
Baseline	9% (8)	4% - 17%	91
Single-Comparison	13% (11)	7% - 21%	87
Double-Comparison	22% (20)	15% - 32%	89
Category-Building	34% (30)	25% - 45%	88

Contrary to prior research, which demonstrated summarization was a less effective way to promote transfer than comparison (Catrambone & Holyoak, 1989; Gentner et al., 2003), the present work found that combining summarization with categorization supports (category-building condition) is an effective way to promote transfer. The category-building led to a higher rate of spontaneous transfer than single comparison, which provides support for the category status hypothesis as a viable account of transfer. The lack of a significant difference between category-building and double comparison provides limited support, since categorization supports in the category-building task led to higher performance than is typically expected of a summarization task (i.e., it was not

significantly lower than double comparison). Further, the lack of differences on the hint-aided transfer assessment suggests that the category-building task's spontaneous transfer advantage cannot be attributed to a differential ability to apply knowledge, but instead results from improved structure-based retrieval. These findings support the category status hypothesis as an alternative account of transfer.

The interpretation of the comparison conditions is less clear. Prior work has demonstrated that single comparison is an effective way to promote transfer (Catrambone & Holyoak, 1989; Gick & Holyoak, 1983), so the lack of an advantage over baseline is puzzling. The present study used a novel stimulus set, and the observed transfer performance is appreciably lower than has been reported with the convergence materials (Catrambone & Holyoak, 1989; Gick & Holyoak, 1983). This stimulus set may be more difficult than commonly used materials, and a single comparison opportunity might require additional support to remain effective under more difficult circumstances.

There was no significant difference between single and double comparison on spontaneous transfer performance, which suggests that an additional comparison opportunity might not appreciably improve abstraction. If the present principle is more difficult than previous materials, an additional comparison opportunity may also be lacking in the support needed to remain effective. While single-comparison did not promote transfer above baseline levels of performance, double-comparison did. This suggests that there may be some small benefit to engaging in the second comparison opportunity.

Experiment 2

Experiment 2 was conducted to conceptually replicate the main findings and clarify some of the outstanding questions of Experiment 1. The same study conditions were used with the exception that a baseline condition was not included. Instead of a problem-solving assessment, a spontaneous reminding task was used. Participants were given a series of cue passages that were superficially distinct from, but contained matching structure with the study materials, and were asked what each cue reminds them of. Since participants were given completed cases as a cue, the entire relational structure of the principle guides the memory search. This is in contrast to the problem-solving transfer assessment that provides only the problem statement as a cue to initiate the memory search. Under less demanding retrieval circumstances, a double-comparison advantage might be accrued. Other modifications were made in an attempt to support the comparison conditions. Both problem-as-a-solution and convergence (Gick & Holyoak, 1983) materials were used to test if the problem-as-a-solution principle was more difficult to retrieve than convergence, since the difficulty of the principle in the first experiment may have been a barrier to comparison success. The instructions for the comparison conditions were modified to increase the symbolic juxtaposition – invitation

to compare through shared labels (Gentner, 2005) – of the cases in another attempt to enhance the comparison task.

The main prediction was that category-building will lead to more structure-based reminders than either comparison condition. Since the procedure was made less demanding in an attempt to promote comparison performance, double-comparison was predicted to have a higher rate of reminding success than single-comparison. Given the overall low rate of transfer in the first experiment, it was predicted that cues for the convergence principle will result in a higher rate of successful reminders than the cues for the problem-as-a-solution principle.

Method

Participants A total of 104 undergraduate students from Binghamton University participated for course credit. Data from three participants were excluded due to a failure to complete the experiment in the allotted time, and another participant was excluded for failing to follow instructions.

Materials and Design The study materials consisted of both the problem-as-a-solution and convergence (Gick & Holyoak, 1983) principles. The study cases for problem-as-a-solution were the same as in Experiment 1. The convergence cases used were as follows: The General, The Commander, Red Adair, and The Fire Chief (Gick & Holyoak, 1983). These cases were rewritten to be comparable in length and grammatical structure to the problem-as-a-solution materials. The order of principles remained constant across participants; problem-as-a-solution occurred first and convergence occurred second.

The reminding assessment consisted of six cue cases. Two cues were used that demonstrated the problem-as-a-solution principle from Experiment 1. The transfer problem used in Experiment 1 was rewritten to include the solution and the other cue involved police infiltrating black markets. The Radiation Problem (Gick & Holyoak, 1983) and The Aquarium (Catrambone & Holyoak, 1989) were rewritten as reminding cues for the convergence principle. Two distractor cases – The Wine Merchant (Gick & Holyoak, 1980) and The Birthday Party (Gick & Holyoak, 1983) – were also included in the reminding assessment in an attempt to disguise the true purpose of the assessment. The order of the cues was constant across participants: distractor, problem-as-a-solution, convergence, distractor, problem-as-a-solution, convergence.

Procedure The study task procedure was similar to the first experiment, with only a few differences. Participants in both comparison conditions received the same task from the first experiment, then repeated it a second time for the convergence materials. In addition, the principles were referred to as separate 'series' to connote that they reflected different principles. In double comparison, the instructions were modified to clearly connote that the first four passages shared important commonalities, and the second four passages also shared important commonalities. The

category-building condition repeated the task for the convergence principle after completion of the task for the problem-as-a-solution principle. The only other difference in the category-building condition was that the category label was replaced with “Conaway Scenario” for the first principle and “Rummel Scenario” for the second principle to clearly reflect the change in principles.

After the study phase all participants were given the reminding packet, which was introduced as a new experiment. Participants were told they would be shown a set of passages, and were supposed to write down anything that each passage reminded them of in as much detail as possible. Participants were then presented with each reminding cue sequentially and made their response. Reminding performance was scored by the first author and an undergraduate research assistant. A successful reminding on the basis of shared structure met at least one of the three following criteria: 1) used the category label or referred to solved problems, 2) referenced one of the cases from the study task, or 3) described the principle from the study task. Any reminders of another cue from within the assessment were considered non-scoring. Both raters agreed on scores for 99.8% of the reminding responses, all disagreements were resolved through discussion.

Results and Discussion

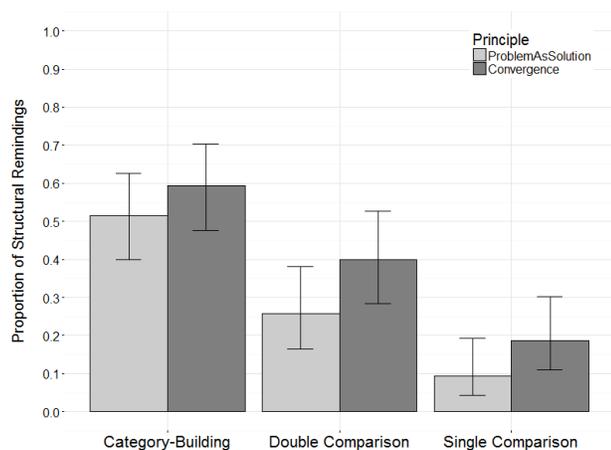


Figure 1: Proportion of structural reminders to target cues by study task and principle. Error bars reflect 95% binomial confidence intervals (Dorai-Raj, 2014).

Reminding performance was modelled trial-wise via a mixed-effects logistic regression (Bates et al., 2015) with the main effect of interest as a predictor and participant included as a random intercept. The predictions concern participants’ responses to only the target cues that had shared structure with the study materials, so only those cues are considered. Category-building led to a higher rate of successful reminders than double-comparison ($\beta = 3.436$, $SE = 1.671$, $Wald Z = 2.057$, $p = .0399$) and single-comparison ($\beta = 6.178$, $SE = 1.649$, $Wald Z = 3.746$, $p < .001$). Double-comparison led to a significantly higher rate of successful reminders than single comparison ($\beta = 2.741$,

$SE = 1.307$, $Wald Z = 2.1$, $p = .036$). Collapsing across condition, convergence cues led to a significantly higher rate of successful reminders than problem-as-a-solution cues ($\beta = 1.5$, $SE = .622$, $Wald Z = 2.41$, $p = .016$) (see Figure 1).

The category-building task led to a higher rate of structurally based reminders to target cues than either comparison condition. This suggests that category-building promotes the spontaneous access and retrieval of relevant structural matches from memory, and that this is driving the spontaneous transfer differences observed in the first experiment. These results provide further support for the category status hypothesis and a successful replication of the main finding in Experiment 1.

These results also address some of the outstanding questions from the first experiment. In contrast to the previous findings, double comparison had a significantly higher rate of structural reminders to target cues than a single comparison opportunity. This suggests that additional comparison opportunities can enhance the retrieval of structurally relevant information from memory. However, we cannot identify which changes were responsible for the observed improvements. Additionally, convergence cues led to a higher rate of successful structural reminders than problem-as-a-solution cues, which may suggest that the convergence materials result in higher rates of transfer than other materials. However, the convergence study materials were always presented after problem-as-a-solution, so future work should explore if this is the result of a practice effect.

General Discussion

When a sequential summarization task was given additional categorization-based supports (the category-building condition), it led to better spontaneous transfer performance than the standard version of the comparison task (single comparison), but did not significantly differ from a task that controlled for case exposure (double comparison). The second experiment replicated the advantage of category-building over single comparison, and found that category-building led to a higher rate of structure-based reminders than double comparison. This supports the conclusion that the spontaneous analogical transfer gains in the category-building condition were due to an increase in retrieval based on matching structure. Taken together, both experiments provide additional support for the category status hypothesis as a viable account of spontaneous transfer.

It is possible that the use of category labels in the category-building condition might confer symbolic juxtaposition (Gentner, 2005), which may allow for abstraction to occur in the absence of the temporal and spatial juxtaposition that is present during comparison. However, this explanation is unlikely the sole factor driving the results. The comparison conditions referred to cases as ‘solved problems’ to control for the use of a category label, and referred to cases from each principle as a coherent ‘series’. The comparison conditions’ controls for the presence of a label likely conferred some degree of

symbolic juxtaposition. If category-building benefits were due to symbolic juxtaposition promoting abstraction, it seems unlikely that a condition which has temporal, spatial, and symbolic juxtaposition (comparison conditions) would perform significantly worse.

The benefits of an additional comparison opportunity are less clear. During the analogical transfer assessment, no advantage of an additional comparison opportunity was accrued. However, in the reminding assessment, an extra comparison opportunity led to a higher rate of retrieval on the basis of shared structure. The benefits of an additional comparison opportunity may occur only under less demanding circumstances, such as being given the full structure as a retrieval cue and not needing to apply the knowledge to solve a problem.

The category status hypothesis can provide an alternative perspective about the success of double-comparison in the second experiment that is not mutually exclusive with the abstraction account. The second instance of comparison provides a chance to build extensional knowledge of the category, since participants are told that the cases are related and participants are given a chance to test hypotheses about why. Further, since some abstraction has likely occurred during the first comparison, the second comparison may afford the opportunity for a bi-directional mapping between generic knowledge of the principle and concrete knowledge of the cases. Future work should further explore the conditions required for additional comparison opportunities to promote spontaneous transfer as well as the role of the category status hypothesis in improving comparison.

There are two possibilities about why category-building and double-comparison led to better retrieval of structural information. First, promoting category status could engage the use of a different type of retrieval process. This retrieval process might occur through the mechanism of categorization as opposed to cue-based reminding. Alternatively, conferring category status might enhance or alter the cue-based reminding process described in Forbus et al. (1995). Future work should try to uncover the mechanism by which these two pathways to improved structure-based retrieval operate.

Acknowledgments

We would like to thank Daria E. Slaby and Ian T. Goodman for their help scoring data, Garrett Honke for creation of the problem-as-a-solution stimulus set, and members of the Learning and Representation in Cognition lab.

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48*, 87-113.

Bates, D., Maechler, M., Bolker, B., Walker, S., (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 1147-1156.

Dorai-Raj, S. (2014). binom: Binomial confidence intervals for several parameterizations. R package version 1.1-1. <https://CRAN.R-project.org/package=binom>

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*, 141-205.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison, (Eds.), *Building object categories in developmental time*. (pp. 245-275). Hillsdale, NJ: Erlbaum.

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. Ahn, R. Goldstone, B. Love, A. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory* (pp.151-175). Washington, DC: American Psychological Association.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*, 393-408.

Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology, 25*, 524-575.

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

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

Goldwater, M. B., & Schalk, L. (2016, March 7). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin, 142*, 729 - 757.

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

Kurtz, K. J., & Honke, G. (2017). *Sorting out the problem of inert knowledge: Using category construction to promote spontaneous transfer*. Manuscript submitted for publication.

Markman, A.B., & Gentner, D. (2002) Structure-mapping in the comparison process. *American Journal of Psychology, 113*, 501-538.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*, 561-574.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 629 – 639.