Title
The Role of Comparison Processes in the Induction of Schemas for Design Styles

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Omata, Takanobu
Holyoak, Keith

Publication Date
2012

Peer reviewed

The Role of Comparison Processes in the Induction of Schemas for Design Styles

Takanobu Omata (takanobu.omata@gmail.com)
Department of Psychology
University of California, Los Angeles
Los Angeles, CA 90095 USA

Keith J. Holyoak (holyoak@lifesci.ucla.edu)
Department of Psychology
University of California, Los Angeles
Los Angeles, CA 90095 USA

Abstract

Considerable evidence supports the effectiveness of close comparison of examples as a means to promote the induction of schemas that support generalization, especially to novel cases that require far transfer. The ease of comparison would appear to be maximized by presenting the to-be-compared cases in close spatial and temporal proximity. However, findings from a number of recent studies have been interpreted as evidence that induction is fostered not by presenting training cases for a single category together (massed practice), but rather by presenting them in an interspersed fashion (spaced practice). We address this apparent paradox in a study in which people are asked to learn the "styles" of furniture designs from a small number of examples of different products (e.g., a bed frame) and then classify examples of entirely different products (e.g., a chandelier). We contrasted a learning procedure based on comparison of examples presented simultaneously with procedures involving processing of individual items, either massed or spaced. Study time was minimized, and generalization was maximized, when learning was based on comparison. In a further study we use structural equation modeling to assess the content of the schemas for visual styles that are acquired by comparison processes. We propose that comparison fosters induction, whereas spacing facilitates retention and retrieval.

Keywords: induction; schema induction; design; spacing effect; structural equation modeling

Introduction

Early work on learning and transfer based on analogy provided strong evidence that close comparison of two examples of a complex category (e.g., problems that can be solved by using converging weak forces) supports subsequent transfer to novel cases that exhibit a similar relational structure (Gick & Holyoak, 1983). The inductive benefit of comparison can arise either as a deliberate learning strategy or as a side effect of applying one solved source problem to an unsolved target problem (Novick & Holyoak, 1991; Ross & Kennedy, 1989). The positive impact of comparison has been demonstrated for both adults (Catrambone & Holyoak, 1989) and young children (Brown, Kane & Echols, 1986; Chen & Daehler, 1989; Holyoak, Junn & Billman, 1984; Kotovsky & Gentner, 1996; Loewenstein & Gentner, 2001; Namy & Gentner, 2002; Star & Rittle-Johnson, 2009). Comparison has been shown to guide schema formation in teaching such complex topics as negotiation strategies (Loewenstein, Thompson, & Gentner, 1999, 2003), and also may play important roles in language learning (Gentner, 2010; Gentner & Namy, 2006). The dominant interpretation of these findings has been that comparison processes foster the induction of a schema for a class of situations, which in turn will facilitate subsequent transfer to additional examples (Gick & Holyoak, 1983).

It would be natural to assume that comparison, and hence induction, will be facilitated by presenting multiple examples simultaneously, or in close temporal proximity. However, this assumption has been challenged. The extensive literature on memory and retention provides robust evidence of an advantage for spaced over massed practice (e.g., Cepeda, Pashler, Vul, Wixted & Rohrer, 2006; Ebbinghaus, 1885/1964). Most of this research has focused on memory for specific items, such as words on a list. However, Kornell and Bjork (2008) showed that the advantage of spaced presentation over massed presentation of training examples extends to a task requiring induction of artistic styles. Classification of new examples was more successful when examples of paintings by different artists were intermixed during training (spaced condition) than when examples of paintings by an individual artist were presented in immediate succession (massed condition). These findings have been interpreted as evidence that spaced presentations actually facilitate participants' generalization. Similar benefits of spacing have been observed in studies of children's category learning (Vlach, Amkowski & Sandhofer, 2012; Vlach, Sandhofer & Kornell, 2008).

On the face of it, the evidence for an advantage of spacing in fostering generalization poses a paradox. If comparison promotes schema induction, and is easier when the examples to be compared are presented in close proximity, it might seem that spaced presentation should hinder rather than help induction; i.e., one might expect spacing to be "the enemy of induction" (E. Z. Rothkopf, quoted by Kornell & Bjork, 2008, p. 585).

However, although simultaneous or massed presentation might be helpful or even necessary for comparison, the mere fact that examples are juxtaposed does not ensure that learners will engage in active comparison. Effective comparison typically is elicited by specific instructions to compare cases and write down commonalities (e.g., Gick & Holyoak, 1983). The benefit of such comparison instructions has been shown to greatly exceed that of simply providing two cases together (even on a single page) without comparison instructions (Gentner, Loewenstein & Thompson, 2003; Thompson, Gentner & Loewenstein, 2000; see also Kurtz, Miao & Gentner, 2001). In order to assess the impact of different presentation conditions, it is therefore important to include a condition in which learners are clearly instructed to perform active comparison of examples.

Accordingly, in Study 1 we directly contrasted a learning procedure based on comparison of examples presented simultaneously with procedures involving processing of individual items presented sequentially, either massed or spaced. We used a novel paradigm in which people attempt to learn realistic styles of furniture and related home décor. After being shown a small number of examples of home décor items of a specific style, participants were asked to judge whether examples of new décor items are of the same style (see Figure 1). This task involves far transfer, since the generalization items included different types of décor items than the training items (e.g., after seeing a dresser, bed frame, fabric and pillowcase set during training, a generalization item required judging the style of a chandelier). Although the relevant cues that might provide the basis for forming a schema are presumably visual, the schema is likely to be quite abstract, and not tied to any single décor type. In Study 2 we employ structural equation modeling in an effort gain insight into the nature of the style schemas acquired via comparison processes.

Study 1

Method

Participants A total of 147 participants (49 in each of three training conditions) were recruited online through Amazon Mechanical Turk (http://www.mturk.com). This system has been demonstrated to produce reliable data in many experimental studies (e.g., Paolacci, Chandler, & Ipeirotis, 2010; Mason, & Suri, 2011). Each participant was paid 60 cents for completing the study, which took about 10 minutes. At the conclusion of the study, information was collected about participants' age, nationality, and possible color blindness. Participants (52 male, 95 female) were all residing in the United States, and ranged in age from 18–61 years (M = 31.5). We tracked IP addresses to ensure that participants were not repeatedly sampled. The participants described above excluded those who failed to complete all of the conditions in the experiment, or reported color blindness.

Materials The materials were 14 color pictures of IKEA furniture and other home décor items (7 each of two styles) printed in the catalog of IKEA 2012 (http://info.ikea-usa.com/Catalog/). The training items for each style (termed Style X and Style Y) consisted of a drawer, a bed frame, fabric, and a duvet set. The tests items used in the subsequent generalization phase were a wall lamp, a (novel) bed frame, and a chandelier (see Figure 1).

Procedure and Design Three conditions, Comparison, Massed and Spaced, were manipulated across participants. In the Comparison condition, the four items of each style, X and Y, were shown together on the screen (randomizing position of items, and counterbalancing order of styles as well as assignment of the labels X and Y to styles). Participants were asked to write down three commonalities for each style. In the Massed and Spaced condition, instructions focused attention on individual items rather than commonalities. In the Massed condition each example of a given style was presented separately, but consecutively; whereas in the Spaced condition examples of the two styles were alternated. In both of the latter conditions participants were asked to write descriptions of each individual item of home décor. Participants advanced through the study phase at their own pace, and their total study time was recorded.

After the study phase, participants were presented with three pairs of new pictures of home décor items. For each pair one was a product of Style X, and the other a product of Style Y (positioned randomly on the left or right). Participants used an 8-point scale to rate which item was from Style X (where a rating of 8 indicated certainty that an item was from Style X).

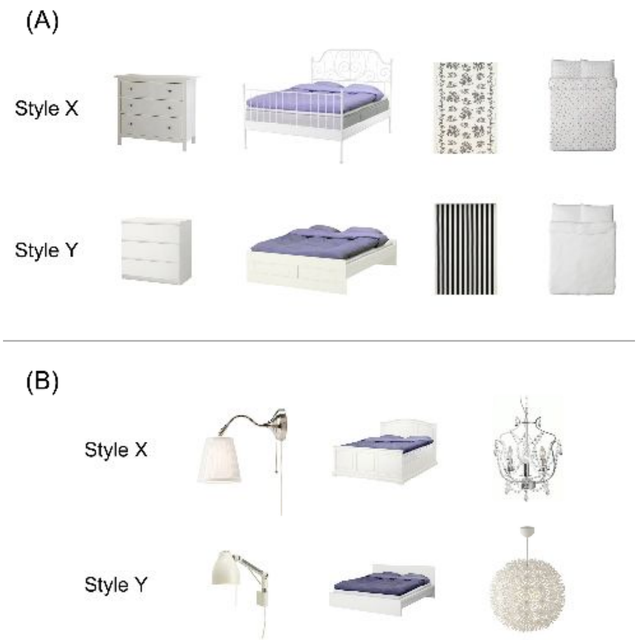

Figure 1. (A). Four examples of each style presented in study phase of Study 1. (B). Three pairs of new pictures presented to the participants in the test phase. (NB: assignment of the labels X and Y was counterbalanced.)

Result and Discussion

Mean study times differed across the three conditions: 8.53 min (Spaced), 7.96 min (Massed), and 4.93 min (Comparison), $F(2, 144) = 16.5$, MSE = 11.2, $p < .0001$. Tukey tests showed that study time was lower for the Comparison condition than either the Massed or Spaced condition ($p < .0001$), whereas the Spaced and Massed conditions did not differ reliably.

Figure 2A presents the mean ratings of the correct item representing Style X on the test phase, plotted such that higher values (max = 8) indicate greater certainty that the correct item was indeed from Style X. A 3 x 3 mixed-factors ANOVA revealed a reliable effect of condition, $F(2, 144) = 8.90$, MSE = 4.72, $p < .001$. There was also a main effect of test item, $F(2, 288) = 3.97$, MSE = 10.66, $p < .05$, but the effect of condition did not vary significantly across items, $F(4, 288) = 1.46$, MSE = 3.91, $p = .22$. Collapsing over items, mean ratings were 5.74 (Spaced), 6.08 (Massed) and 6.79 (Comparison). Tukey tests revealed that generalization was more accurate for the Comparison condition than either the Spaced ($p < .01$) or Massed condition ($p < .05$), whereas the Spaced and Massed conditions did not reliably differ.

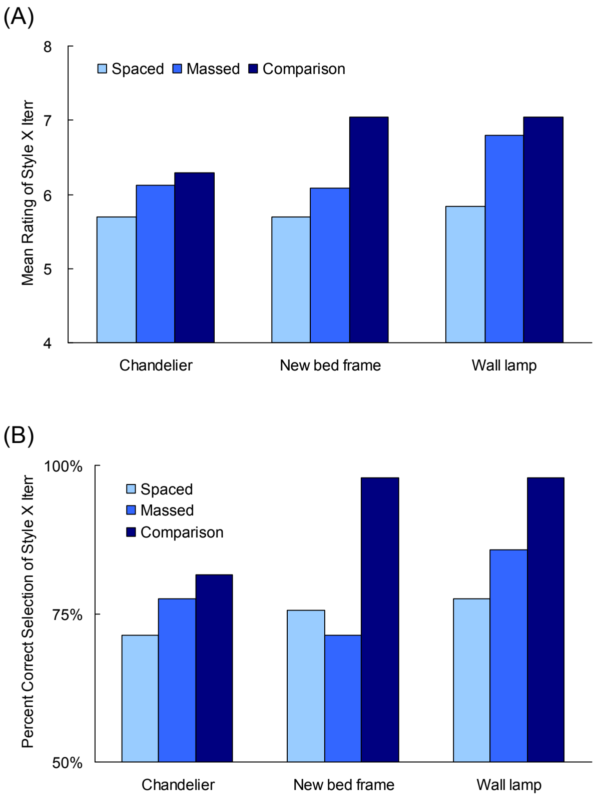

Figure 2. (A). Mean ratings of correct item from Style X for three pairs of new home décor items presented in test phase of Study 1. (B). Mean percentage of correctly selecting item from Style X for each of the three pairs.

We also calculated the percentage of correct choices by reducing the 8-point rating scale to a binary decision, breaking it at the midpoint (ratings of 1-4 vs. 5-8). As shown in Figure 2B, Comparison instructions yielded higher overall accuracy (93%) than either the Massed (78%) or Spaced (75%) condition. After aggregating across the three items, a one-way between-subjects ANOVA revealed a reliable effect of condition, $F(2, 146) = 7.11$, MSE = 0.86, $p < .01$. Tukey tests confirmed that generalization was more accurate for the Comparison condition than either the Spaced ($p < .01$) or Massed condition ($p < .05$), whereas the latter two conditions did not reliably differ. The results of Study 1 thus provide clear evidence that comparison instructions designed to encourage induction of schemas for styles of home décor support more efficient learning and generalization than do presentation methods that focus attention on individual examples. The Comparison condition required less study time than either the Spaced or Massed conditions, yet yielded more successful generalization.

Study 2

The results of Study 1 support the hypothesis that comparison instructions aid in inducing a schema for styles of furniture design, facilitating subsequent generalization. In Study 2 we examined the possible causal relationships between content of a style schema and generalization, using structural equation modeling to infer relevant latent variables. For modeling purposes only a single condition was run, with study, test, and evaluation phases. Participants first studied examples of multiple styles of home décor items, focusing on one particular style (called Style X), then judged which of new examples were from Style X, and finally rated descriptive words as to how well they fit their impression of Style X. The ratings of descriptors were used to estimate latent variables representing a schema for Style X, with the aim of predicting the generalization data from the test phase.

Method

Participants A total of 175 participants (50 male, 125 female), ranging in age from 18–70 years ($M = 32.4$), all residing in the United States, were recruited through Amazon Mechanical Turk. Participants were paid 60 cents for participating in the experiment, which took about 10-15 minutes.

Materials The materials were 23 pictures of IKEA home décor items, five representing each of five styles (labeled X, A, B, C, and D), selected from the same source as in Study 1. Style X (a specific style that was constant for all participants) was the focus of our modeling effort (see Figures 3-4). In addition, a list of 30 descriptive words was created for use in the evaluation phase (see Table 1). To create this list, 118 participants in an initial survey (also conducted on Mechanical Turk) were asked to generate words describing Style X (based on the same three examples used in the study phase of Study 2). Using these generated words as a starting point, we selected relatively high frequency and non-redundant words, and added several additional words (as distractors) that had *not* been generated to describe Style X.

Procedure and Design In the study phase, three pictures (a drawer, a bed frame, and a chandelier) were presented together for each style (see Figure 3). The study phase began with three steps that applied only to the examples of Style X (the style that was the direct focus of structural equation modeling). Participants (1) listed three common features of the examples of Style X; (2) described their "impression, feel, or sense of Style X" based on its common features; and (3) described the "personality" that Style X would represent if it were thought of as a person. Next, participants saw all five examples of each type of home décor item (e.g., stools of each of the Styles X and A-D), and were asked to compare them and describe the difference between Style X and the other styles. This task was repeated for each of the three types of training items. The presentation order of pictures in Style X and the other styles was fully randomized.

During the subsequent test phase (see Figure 4), pairs of examples of four novel types of home décor items were presented (stools, wall lamps, duvet sets, and fabric). Each pair included one example of Style X and one example of some other style (Styles A-D). For each pair, participants rated which example was more likely to be from Style X, using the same 8-point scale as had been used in Study 1.

Finally, in the evaluation phase, participants were presented with the 30 descriptors in random order. Participants used a 4-point rating scale to evaluate how well word captured their impression of Style X (with a rating of 4 indicating that the descriptor "definitely" fit Style X).

Figure 3. Three examples of each style presented in study phase of Study 2.

Figure 4. Four pairs of new pictures presented to the participants in the test phase of Study 2. For each pair, the item in the left column is from Style X; that on the right is the foil, drawn from Styles A-D.

Table 1. Set of 30 descriptors used in evaluation phase of Study 2.

Words produced as descriptors of Style X
feminine (25), old-fashioned (18), girly (17), traditional (9), plain (9), fancy (8), country (7), elegant (6), stylish (4), pretty (4), conservative (3), familiar (3), light (3), luxury (3), relax (3), cheery/active (2), romantic (2), soft (2), warm (2), modern (1), unisex (1)

NB: number of respondents in initial survey (out of 118) who produced each descriptor in response to Style X is indicated in parentheses.

Distractors (not produced in response to Style X)
casual, cool, fashionable, formal, gorgeous, hard, masculine, natural, wild

Results and Discussion

Prior to conducting structural equation modeling, we checked descriptive statistics from the evaluation phase, identifying five descriptors with means ratings above 2.50 on the 4-point scale of applicability to Style X. We then performed exploratory factor analysis using these five scales, obtaining a promax rotation by maximum likelihood estimation using SPSS. Contribution ratios and eigenvalues were computed, and two factors were retained (eigenvalues > 1.0), which explained 60.0% of total variance. The inter-factor correlation between the two factors was 0.24.

We then performed structural equation modeling using EQS version 6.1 (http://www.mvsoft.com/) in order to identify apparent causal relationships between a style schema (based on the latent variables derived from ratings of descriptors) and the ability to discriminate examples of Style X on the test phase. The analysis identified a model with three latent factors that provided an excellent statistical fit to the data, χ^2 (24) = 31.3, *p* =.14. The value of Root Mean Square Error of Approximation (RMSEA) was 0.04, and the value of the Comparative Fit Index (CFI) was 0.98.

The structural equation model is displayed in Figure 5. The factor labeled D (Discriminability) summarizes the ability to discriminate Style X from others on the test phase, based on the four items used to assess generalization of the knowledge about Style X acquired during the study phase. The D factor was in turn predicted by factors F (Femininity) and C (Classic), derived from the ratings of descriptors during the evaluation phase. Specifically, factor F was associated with ratings of the words "feminine", "fancy", and "girly", while factor C was associated with ratings of "traditional" and "old-fashioned". The factor loadings for Factors F and C exceeded 0.50 and for factor D were over 0.20. Factors F and C had positive path coefficients between each other as well as to factor D. Intuitively, participants who associated Style X with descriptors indicative of a feminine and classic image were best able to discriminate examples of Style X from alternatives on the generalization test.

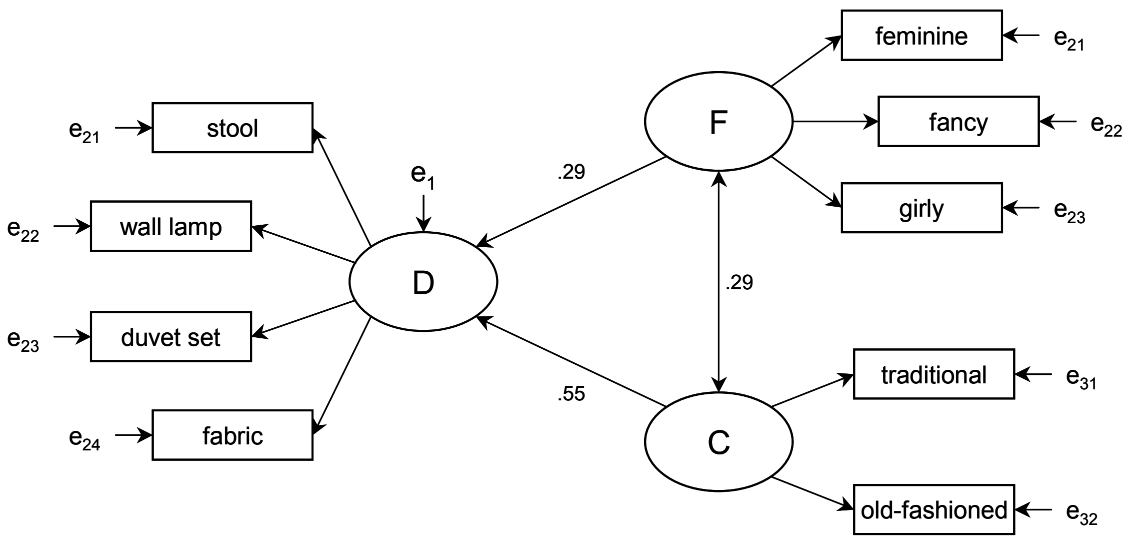

Figure 5. The model for predicting generalization of Style X to novel items on the test phase, derived by structural equation modeling in Study 2. Values of correlation coefficients are indicated. Error terms are denoted by subscripted *e*.

General Discussion

Using realistic visual categories (styles of furniture and other home décor items) we showed in Study 1 that active comparison of examples of a category is more efficient (reduced study time and greater accuracy) in promoting subsequent generalization than is processing of individual examples, whether massed or spaced. These findings support the hypothesis that schema induction underlies successful generalization of a visual style. In Study 2 we attempted to assess the content of a design schema. Using structural equation modeling, we showed that after learning by comparison, generalization performance can be predicted by latent variables corresponding to relatively abstract concepts derived from descriptive terms.

Taken together, these findings support the usefulness of the concept of schema induction as a basis for generalization with complex categories. The results of Study 1, in particular, suggest that it would be a mistake to view spaced presentation *per se* as an optimal procedure for promoting induction and generalization. Not only did the Spaced condition yield poorer generalization (despite longer study time) than the Comparison condition, but it showed no advantage (trend in wrong direction) relative to the Massed condition. Massed presentation, although not sufficient to reliably trigger comparison processes, may nonetheless make such processing more likely.

Based on other findings in the literature (e.g., Kornell & Bjork, 2008), there clearly are some conditions under which spacing improves later generalization performance, particularly in comparison to massed presentations that focus on encoding of individual items. In fact, comparison and spacing may well convey distinct and complementary benefits for categorization performance. In terms of the traditional stages of memory—encoding, retention and retrieval—schema induction can be viewed as a special type of encoding that focuses on abstraction of general cues to category membership, as opposed to encoding of individual training examples. We suggest that comparison facilitates schema induction during encoding, whereas spacing has its greatest impact on retention and retrieval. For complex categories of the sort used in the present studies, spacing likely makes schema induction more difficult; however, spacing may sometimes provide compensatory benefits in increasing retention and later retrieval of knowledge. This hypothesis suggests that generalization might be optimized by early comparison of concurrently-presented examples, followed by subsequent spaced presentation of additional examples. Further research will be needed to explore the potential interactions between factors that guide induction of schemas and those that aid their retention and retrieval.

Acknowledgments

Preparation of this paper was supported by a gift from Canon Incorporated, and by grant N000140810186 from the Office of Naval Research. We thank Airom Bleicher for assistance in running the studies.

References

Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development*, *1*, 103-121.

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

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132,* 354–380.

Chen, Z., & Daehler, M. W. (1989). Positive and negative transfer in analogical problem solving by 6-year-old children. *Cognitive Development*, *4*, 327-344.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger, C. E. Bussenius, & E. R. Hilgard, Trans.). New York: Dover. (Original work published 1885.)

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

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

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

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., Junn, E. N., & Billman, D. O. (1984). Development of analogical problem-solving skill. *Child Development, 55*, 2042-2055.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science, 19,* 585–592.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*, 2797-2822.

Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences*, *10*, 417–446.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development, 2*, 189-219.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin and Review*, *6*, 586-597.

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.

Mason, W.A., & Suri, S. (2011). How to use Mechanical Turk for cognitive science research. In L. Carlson, C. Hőlscher, & T. F. Shipley (Eds.), *Proceedings of the 33nd Annual Conference of the Cognitive Science Society* (pp. 66-67). Austin, TX: Cognitive Science Society.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*, 5-15 10.

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 398-415.

Paolacci, G., Chandler, J., & Ipeirotis, P.G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*, 411-419.

Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 42-55.

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102,* 408–426.

Thompson, L., Gentner, D., & Loewenstein, J. (2000). Avoiding missed opportunities in managerial life: Analogical training more powerful than individual case training. *Organization Behavior and Human Decision Processes*, *82*, 60-75.

Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). Same time or part in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 246-254.

Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition*, *109*, 163-167.