UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Comparison Promotes the Spontaneous Transfer of Relational Categories

Permalink

https://escholarship.org/uc/item/0np062mv

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

Authors

Snoddy, Sean

Kurtz, Kenneth

Publication Date

2021

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Comparison Promotes the Spontaneous Transfer of Relational Categories

Sean Snoddy (ssnoddy1@binghamton.edu)
Kenneth J. Kurtz (kkurtz@binghamton.edu)

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

Abstract

Snoddy and Kurtz (2020) demonstrated spontaneous transfer of relational categories to new learning. Recognition memory data suggested that transfer was driven by schematization during learning. In the present study, we explored whether schema abstraction underlies transfer and recognition effects. Participants were assigned to condition based on the type of initial learning: classification with comparison, supervised observation with comparison, single-item supervised observation, or baseline (no learning). After initial learning, participants underwent a study phase and recognition test on novel stimuli followed by a target category learning task involving the same underlying category structures expressed in a new domain. During the recognition test, all conditions led to increased false alarms relative to baseline. Only the comparison conditions exhibited analogical transfer on the target category learning task. Results suggest that comparison facilitates the transfer of relational categories (due to schema abstraction), but recognition memory effects may be driven by more general categorization mechanisms.

Keywords: category learning; relational categories; comparison; analogical transfer; recognition memory

Introduction

Categorization enables us to organize knowledge, summarize like instances, and generalize knowledge to novel situations. The extant category learning literature has chiefly focused on attribute categories (e.g., *bird*), which are characterized by similarity in members' intrinsic features (i.e., *birds* have wings, feathers, etc.; Gentner & Kurtz, 2005). In contrast, relational categories capture extrinsic similarity in relational structure (i.e., turning on a lamp and melting ice both reflect *dispersion* of energy) between members (Gentner & Kurtz, 2005). As structural similarity is the basis of relational categories, they are not confined to a single domain like attribute categories (Gentner & Kurtz, 2005) and provide a way to study cross-domain generalization of concepts.

Cross-domain generalization of relational concepts is similar, if not identical, to analogical transfer—applying knowledge of a relational principle (e.g., natural selection) acquired in one domain (e.g., biology) to solve a problem in another (e.g., business). The canonical analogical transfer effect is that when the problem-solving task is administered as a separate, unrelated experiment (i.e., spontaneously), transfer is notoriously difficult to achieve, however, transfer readily occurs when participants are given a hint about relevant prior knowledge (Gick & Holyoak, 1980). The ease of hint-aided transfer demonstrates that the chief difficulty in spontaneous transfer stems from accessing germane, albeit superficially dissimilar knowledge.

A key finding in the transfer literature is that comparison of multiple analogs during study improves spontaneous transfer (Gick & Holyoak, 1983). The benefits of comparison arise from a structural alignment process that supports the formation of abstract schemas (i.e., abstract representations of relational structure) by filtering out surface-level mismatches such that the core relational structure remains (Gick & Holyoak, 1983; Markman & Gentner, 2000). These abstract schemas are more accessible in memory because they lack superficial mismatches with retrieval cues that are prohibitive to spontaneous access (Gentner, Rattermann, & Forbus, 1993). The primary goal of the present work is to test whether comparison is able to support spontaneous transfer via schema abstraction in relational category learning.

Prior work has begun to explore the effect of comparison in relational category learning. Kurtz, Boukrina, and Gentner (2013) demonstrated that comparison-based classification led to improved cued, cross-domain transfer relative to single-item classification. In related work, Patterson and Kurtz (2020) found a comparison advantage over single-item presentation under a supervised observational learning mode (i.e., labelled pairs), but not with a guess-and-correct classification task. These studies suggest that comparison is able to support transfer of relational categories. However, as both studies cued relevant knowledge by employing the same category labels in the training and transfer phases, their results most closely align with hint-aided transfer in Gick and Holyoak (1980, 1983). As hint-aided transfer does not involve spontaneously accessing knowledge, it remains unclear whether comparison-based category learning leads to improved transfer through better encoding of the category structures or if it can also encourage spontaneous access.

Snoddy and Kurtz (2020) explored whether category learning tasks are capable of supporting spontaneous access and subsequent transfer. Participants who underwent single-item classification to acquire relational categories during a base category learning phase (BCL) demonstrated larger improvements across training in a target category learning phase (TCL)—where the same underlying category structures from BCL were presented in a new domain with novel category labels—relative to participants learning the category structures for the first time. These results suggested that participants were leveraging knowledge of the category structures acquired during BCL to facilitate performance on TCL; thus, providing evidence of relational categories being spontaneously transferred across domains. We sought to

extend Snoddy and Kurtz (2020) by testing whether comparison-based relational category learning can further improve spontaneous transfer by better supporting access.

If comparison effectively encourages access in relational category learning, such effects should be evidenced in memory measures as categorization and recognition memory are thought to be driven by a shared system (Nosofsky & Zaki, 1998). De Brigard, Brady, Ruzic, and Schacter (2017) found that learning about an attribute category led to an increase in both hits (i.e., correct recognition of studied items) and false alarms (i.e., spurious recognition of novel items as being previously studied) for novel category-consistent items relative to non-learned items on a recognition memory test. De Brigard et al. (2017) explained these effects in terms of a schematization process—category learning encourages a focus on diagnostic information, which in turn leads to novel recognition items being treated as if they had been experienced during learning.

For relational categories, focusing on diagnostic information involves attending to structural information while discounting the superficial information (i.e., schema abstraction). To extend De Brigard et al. (2017) to relational categories, Snoddy and Kurtz (2020) included study and recognition test phases between BCL and TCL. Their results provided suggestive evidence for a schematization account—category learning led to an increase in false alarms, but not hits. This suggested that category learning encouraged abstraction, which provided a familiarity-based route to hits and false alarms, whereas the baseline led to a superficial bias that supports only hits (Yonelinas, 2002). A comparison-based task, which better supports schema abstraction, should result in a larger false alarm rate than a single-item task, which would provide convergent evidence for this account.

The motivation of the present study is to explore the role of comparison-based relational category learning on both spontaneous transfer and recognition memory. To this end, the design was based on that of Snoddy and Kurtz (2020). BCL training provided an opportunity to acquire relational categories; it included comparison-based tasks, a single-item supervised observation task (OBS-SINGLE), and a no learning Baseline task. Due to the results of Patterson and Kurtz (2020), both classification-based (CLASS-COMP) and observation-based (OBS-COMP) comparison tasks were included. Observation tasks do not provide a measure of training accuracy, so a test was included at the end of BCL to assess mastery of the category structures.

After BCL, participants underwent a recognition memory phase to test the schematization account. Participants were then told that they were moving on to a new experiment, and the TCL phase was administered in a new stimulus domain to serve as an opportunity to spontaneously transfer knowledge of the category structures acquired during BCL. Due to differences between BCL tasks, a condition-neutral task was used during TCL: yes/no endorsement of whether an item-label pairing is correct (*cf.* Patterson & Kurtz, 2020). We also sought to explore the effect of providing a hint on transfer success. Two different hints were used to test whether initial learning in the transfer domain alters the hint's effectiveness. A second observation comparison task (OBS-COMP-HINT) was given a hint at the onset of TCL, while the remaining tasks received the hint halfway through TCL training.

If comparison effectively encourages schema abstraction, then comparison-based category learning tasks will lead to larger transfer effects during TCL learning than both OBS-SINGLE and Baseline. If the recognition effects are driven by a process akin to schema abstraction, then comparison tasks should lead to an increased rate of false alarms during the recognition test. The inclusion of OBS-SINGLE and Baseline allows for an extension of Snoddy and Kurtz (2020) to a more difficult transfer assessment (i.e., transfer to a new task). Comparison of spontaneous and hint-aided transfer was exploratory.

Method

Participants

A total of 283 undergraduate students from Binghamton University participated. Participants were randomly assigned to Baseline (N = 57), CLASS-COMP (N= 60), OBS-COMP (N = 56), OBS-COMP-HINT (N = 64), or OBS-SINGLE (N = 56) conditions.

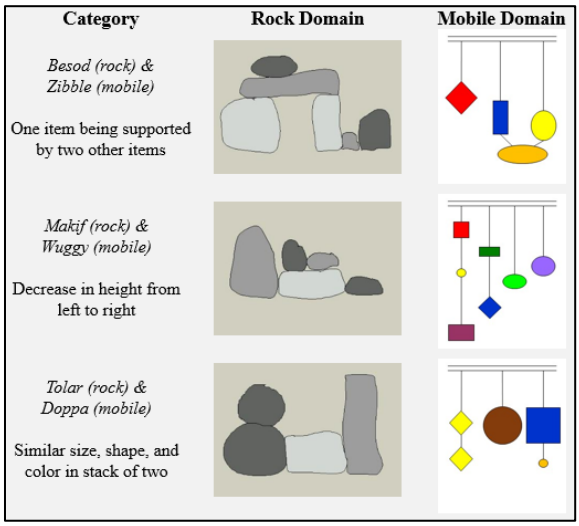


Figure 1: A description of each of the three category structures, category labels, and sample stimuli.

Materials and Design

The stimuli consisted of 84 images in the domains of rock arrangements and mobile-like displays adapted from previous work (Kurtz et al., 2013; Patterson & Kurtz, 2020; Snoddy & Kurtz, 2020). Each stimulus instantiated one of three mutually exclusive category structures: 1) support, 2) monotonicity, or 3) symmetry (for a description the category structures and example items from each domain, see Figure 1). Both superficial attributes of individual elements (rocks or geometric shapes) within each item as well as the number

of intra-item elements were varied and not predictive of category membership. The Mobile domain was associated with a distinct set of category labels ("Zibble", "Wuggy", and "Doppa") from the Rock domain ("Besod", "Makif", and "Tolar") and inverted the same category structures over the x-axis. Following from previous work (Kurtz et al., 2013; Patterson & Kurtz, 2020), the Rock domain was used for training (BCL and recognition memory) whereas the Mobile domain was used for the transfer assessment (TCL).

For each participant, the 48 items in the Rock domain were randomly assigned to either BCL (24 items) or recognition phases (24 items); there was no item overlap between phases. The pseudo-random assignment of items to phases ensured that all categories were always equally represented. Twelve BCL items were randomly selected to be presented during the BCL test. Recognition items were randomly assigned to the study set and used as old items on the test (12 items) or to be used as novel lures at test (12 items). To control for the number of training items between phases, a randomly selected subset of 24 Mobile items (out of 36 possible) were presented during TCL.

Procedure

Due to Covid-19 restrictions, participants were recruited through the university subject pool for a web-based experiment conducted through a JavaScript program on participants' computers (phones/tablets were prohibited). To create the appearance that BCL/recognition and TCL phases were separate experiments, this study was embedded within a larger experimental session (none of the other experiments involved analogy or relational categories). The phases of this experiment were presented in a fixed order (BCL/recognition followed by TCL); the order of the overall experimental session was randomized across participants. (For an overview of the experimental design and procedure, see Table 1).

Table 1: Summary of the Experimental Design and Procedure

	BCL Task			
Experimental Phase	**Baseline**	**OBS-SINGLE**	**OBS-COMP & OBS-COMP-HINT**	**CLASS-COMP**
1.1 BCL Training (Rock Domain)	--	Observation & Single-Item Presentation	Observation & Same-Class Comparison	Classification & Same-Class Comparison
1.2 BCL Test (Rock Domain)	--	Single-Item Classification (no feedback)		
2.1 Recognition Study (Rock Domain)	Same for All Conditions (separate items from BCL)			
2.2 Recognition Test (Rock Domain)	Same for All Conditions (separate items from BCL)			
"Next Experiment" Transition				
3. TCL Training (Transfer to Mobile Domain)	No Hint	Block 4 Hint	Block 4 Hint (OBS-COMP) Block 1 Hint (OBS-COMP-HINT)	Block 4 Hint

BCL Phase Participants in the category learning conditions began the experiment with the BCL phase (except Baseline, which omitted the BCL phase). The experiment instructions began with a cover story about rock arrangements made by the (fictious) "Ladua culture." Participants were tasked with trying to figure out what makes a given arrangement belong to a specific "type of rock arrangement" and told they would later be tested on their knowledge. Task-specific instructions were appended to the general instructions that identified: 1) whether participants would be presented with a single item on each trial (OBS-SINGLE) or a pair of items from the same category (comparison conditions), and 2) whether the correct type would be shown for each item (observation conditions) or participants would make a decision about the correct type and receive corrective feedback (CLASS-COMP).

For the OBS-SINGLE condition, a single stimulus was presented in the center of the screen at the start of each trial. After 500 MS, the correct category label was presented below each item (e.g., "This is a Besod"). Participants were prompted to engage in self-paced study of the item-label pairing and advanced to the next trial by clicking a button located at the bottom of the screen. This task was repeated across three blocks of training. Within each block, all 24 items were sequentially presented in a randomized order.

Participants in the OBS-COMP and OBS-COMP-HINT conditions received a similar task as the OBS-SINGLE condition. The critical differences were that a pair of same-category stimuli were presented side-by-side in the center of the screen and the presentation of the category label was modified to reflect the pair of items (e.g., "These are both a Besod"). Participants engaged in self-paced study of the pairs of same category-items across three blocks of training. To control for item exposure, each block of the comparison tasks consisted of 12 sequentially presented same-category pairs (Kurtz et al., 2013; Patterson & Kurtz, 2020). Items were randomly assigned to pairs at the start of each block.

In the CLASS-COMP condition, a pair of same-category stimuli were presented side-by-side in the center of the screen on each trial. After 500 MS, participants were told that both items belonged to the same category and queried to provide a classification response by clicking on one of the three displayed response buttons containing a category label. Corrective feedback was provided after each response ("Correct! Both of these are a Besod" or "Incorrect… Both of these are a Besod"). Participants completed three blocks of training for the 12 pairs of same-category items.

After BCL training, the BCL single-item classification test was administered. Participants were presented with instructions stating that they would be tested on their knowledge of the category structures from the training phase. For each of the 12 test items, a single stimulus was presented in the center of the screen. After 500 MS, participants were queried to provide a classification response by clicking on one of the three displayed response buttons containing a category label. No corrective feedback was provided.

Recognition Memory Phase Participants then proceeded to the recognition memory phase (Baseline began the experiment here). During the study portion of this phase, participants received instructions stating that they were now in the "study phase", they would see a series of new items, and their task was to study each item so that they could recognize it as belonging to the "study set" later. Participants

were then sequentially presented with 12 items for self-paced study. Each item was centered on the computer screen with a "next item" button below it that participants clicked when ready to advance. An inter-trial interval of 500 MS was used for both study and recognition test. No information about category membership was provided in either task.

The recognition test was administered immediately after study. Instructions stated that participants would be shown a series of items and their task was to decide whether each item was presented in the "study set". All 24 recognition items (12 old items from study and 12 novel lures) were sequentially presented in a randomized order. Upon presentation of each item, participants were prompted to decide whether the item was from the study set via a "yes" or "no" judgement. No corrective feedback was provided. At the end of the recognition test, participants were thanked and told they would be advancing to the next experiment.

TCL Phase All participants were told that they were beginning a new experiment and proceeded to the TCL phase. To this end, the same changes used to distinguish between experiments within the session were applied to the BCL and TCL phases—TCL started with a general, transitional experiment instruction screen and a new color scheme and font was used during TCL. Further changes were made to induce this context shift: 1) task instructions were presented in generic terms instead of a cover story, 2) items were from the Mobile domain, and 3) a new set of category labels was used. This context shift provides limited retrieval cues—aside from structural similarity of BCL and TCL items—and creates an opportunity to transfer knowledge of the category structures from BCL to facilitate learning during TCL.

Due to differences in BCL tasks, TCL used a condition-neutral endorsement task. Participants were sequentially presented with each of the 24 items from the Mobile domain in a randomized order. Items were pseudo-randomly paired with a label such that half of the items were paired with the correct category label. On each trial, participants were asked to judge whether the item-label pairing was correct ("Is this a Doppa?") by clicking response buttons labeled "yes" and "no." Corrective feedback was provided after each response.

TCL was divided into spontaneous and hint-aided phases. During spontaneous transfer, participants were presented with a generic version of the instructions from BCL. OBS-SINGLE, OBS-COMP, and CLASS-COMP conditions then underwent 3 blocks of endorsement training for all 24 items. After completing the third block of training, participants were shown a second instruction screen that contained a hint, which stated that the categories from BCL will be helpful in learning the present ones. Participants in these conditions then underwent an additional 3 blocks of hint-aided endorsement training, where a reminder of the hint ("*Hint*: think about the categories of rock arrangements") was appended to each trial's instructions. In OBS-COMP-HINT, participants received the same hint-aided instructions at the start of the first block of TCL followed by six blocks of hint-aided endorsement training. Participants in the Baseline condition did not receive hint-aided instructions (as there is no relevant prior knowledge to cue) and underwent 6 blocks of spontaneous endorsement training.

Results

All analyses used mixed-effects logistic regression models that predicted trial-wise performance. Specific dependent measures and predictor variables are described for each analysis. A random effect structure that allowed both participant and item to vary as random intercepts accounted for the most variance while still allowing models to converge.

BCL Test

There were no *a priori* predictions regarding BCL test performance, but any differences between conditions would suggest that test performance needs to be controlled for in subsequent analyses. The mixed-effects model predicted trial-wise accuracy with BCL task. There were no significant differences between conditions (all *p*s > .444; see Figure 2). This suggests that all training tasks led to comparable levels of mastery of the category structures.

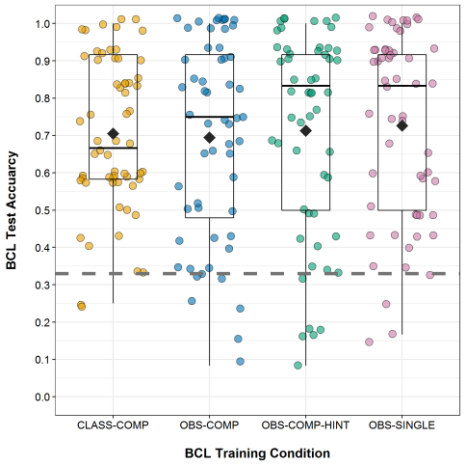


Figure 2: BCL test accuracy by BCL training task. Points reflect individual participants' unadjusted mean accuracy, boxplots reflect unadjusted descriptive statistics, diamonds reflect adjusted mean accuracy, and the dashed line reflects chance performance.

Recognition Test

To test the predictions of the schematization account—more false alarms in comparison conditions than OBS-SINGLE, more false alarms in OBS-SINGLE than Baseline, and an exploratory analysis of hit rates—two separate mixed effects logistic regression models that predicted the outcome of interest (false alarm or hit rates) with BCL task were used. The first analysis was for the novel lures not presented during the study phase (Figure 3, left panel). In support of our prediction, the Baseline condition led to significantly fewer false alarms than CLASS-COMP (β = 0.707, *SE* = 0.206,

Wald Z = 3.434, $p < .001$), OBS-COMP ($\beta = 0.499$, *SE* = 0.209, *Wald Z* = 2.383, $p = .017$), OBS-COMP-HINT ($\beta = 0.498$, *SE* = 0.212, *Wald Z* = 2.348, $p = .019$), and OBS-SINGLE ($\beta = 0.641$, *SE* = 0.209, *Wald Z* = 3.069, $p = .002$). In contrast to our prediction, OBS-SINGLE did not significantly differ from any of the comparison conditions (all *p*s > .489). There were no significant differences between the comparison conditions (all *p*s > .304). Similar to Snoddy and Kurtz (2020), these results demonstrated that prior category learning (regardless of the category learning task) increased the likelihood of false alarms. However, providing increased support for schema abstraction in the comparison conditions did not appear to further increase false alarm rate.

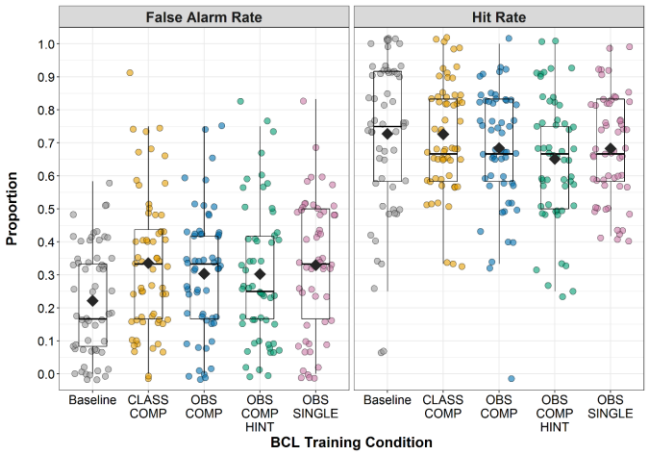


Figure 3: Proportion of false alarms and hits by BCL condition. Points reflect individual subjects' unadjusted proportion and boxplots reflect unadjusted descriptive statistics. Diamonds reflect adjusted proportions.

With respect to the analysis of old items presented during the study phase (Figure 3, right panel), the Baseline condition led to a significantly larger hit rate than OBS-COMP-HINT ($\beta = 0.41$, *SE* = 0.184, *Wald Z* = 2.223, $p = .026$), but did not significantly differ from the other conditions (all *p*s > .155). OBS-SINGLE did not significantly differ from any of the comparison conditions (all *p*s > .217). For the comparison conditions, only CLASS-COMP led to a significantly higher hit rate than OBS-COMP-HINT ($\beta = 0.373$, *SE* = 0.181, *Wald Z* = 2.063, $p = .039$); there were no other significant differences between the comparison conditions (all *p*s > .213). As in Snoddy and Kurtz (2020), prior category learning did not produce a consistent increase in hit rate relative to Baseline. Overall, these findings provided a successful replication of Snoddy and Kurtz (2020)—prior BCL leads to an increase in false alarm rate but not hit rate, regardless of the BCL task, relative to a no learning baseline condition.

TCL Training (Transfer)

Separate models were applied to spontaneous endorsement (blocks 1-3) and hint-aided endorsement (blocks 4-6) data. Each model predicted trial-wise endorsement accuracy with BCL task, block, and their interaction.

Spontaneous Transfer The primary prediction for spontaneous transfer was that all category learning tasks would lead to improved endorsement accuracy relative to Baseline. In support of this prediction, the Baseline condition was associated with significantly lower endorsement accuracy than CLASS-COMP ($\beta = -0.5$, *SE* = 0.187, *Wald Z* = -2.668, $p = .008$) and OBS-COMP ($\beta = -0.406$, *SE* = 0.189, *Wald Z* = -2.146, $p = .032$) conditions. In contrast, there was no significant difference between Baseline and the OBS-SINGLE condition ($\beta = 0.259$, *SE* = 0.187, *Wald Z* = 1.382, $p = .167$). There was a significant effect of block in all conditions—Baseline ($\beta = 0.459$, *SE* = 0.044, *Wald Z* = 10.331, $p < .001$), OBS-SINGLE ($\beta = 0.222$, *SE* = 0.043, *Wald Z* = 5.148, $p < .001$), OBS-COMP ($\beta = 0.301$, *SE* = 0.045, *Wald Z* = 6.62, $p < .001$), and CLASS-COMP ($\beta = 0.28$, *SE* = 0.045, *Wald Z* = 6.239, $p < .001$)—where endorsement accuracy increased across blocks of training. Significant simple interaction terms qualified these effects. There were larger improvements across endorsement training in the Baseline condition relative to CLASS-COMP ($\beta = 0.179$, *SE* = 0.063, *Wald Z* = 2.841, $p = .005$), OBS-COMP ($\beta = 0.158$, *SE* = 0.064, *Wald Z* = 2.49, $p = .013$), and OBS-SINGLE ($\beta = 0.238$, *SE* = 0.062, *Wald Z* = 3.843, $p < .001$) conditions (Figure 4). These results provide evidence of transfer in the comparison conditions, but not OBS-SINGLE, and demonstrate that transfer occurred in the early blocks of training. As OBS-SINGLE did not lead to a transfer effect, it suggests that the single-item classification transfer effect observed by Snoddy and Kurtz (2020) does not extend to a more difficult transfer assessment without the additional support conferred by comparison.

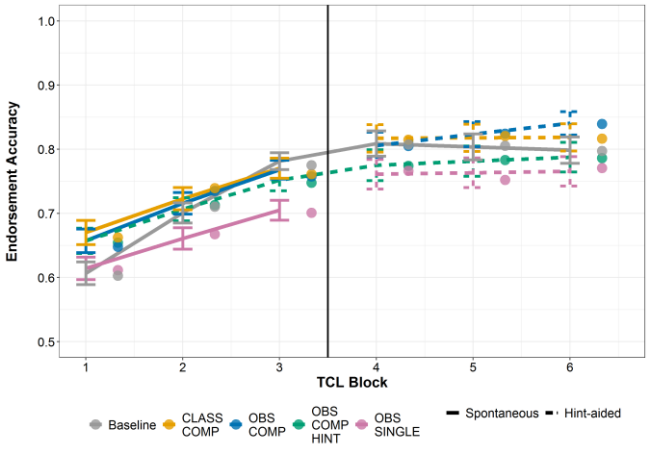


Figure 4: Lines reflect adjusted mean TCL endorsement accuracy across blocks of training for each BCL task. Right offset points are unadjusted means. Error bars are ± 1 SEM. Solid lines reflect spontaneous transfer instructions and dashed lines reflect hint-aided instructions. The vertical black line reflects the introduction of hint-aided instructions.

A second spontaneous transfer prediction was greater TCL accuracy in the comparison conditions than OBS-SINGLE. In contrast to this prediction, there were no significant differences between OBS-SINGLE and the comparison conditions (all *p*s > .199) nor significant simple interactions (all *p*s > .204). While exploratory, there was no significant difference (β = -0.094, *SE* = 0.19, *Wald Z* = -0.496, *p* = .62) nor interaction (β = -0.021, *SE* = 0.064, *Wald Z* = -0.33, *p* = .741) between OBS-COMP and CLASS-COMP. In contrast to the cued-transfer results of Patterson and Kurtz (2020), both observational and classification comparison tasks appeared to be comparably effective in supporting spontaneous transfer.

Hint-aided Transfer The first hint-aided transfer analysis explored the effect of providing a hint at the onset of TCL by contrasting the first three blocks of OBS-COMP-HINT with the other conditions undergoing spontaneous transfer. Performance in OBS-COMP-HINT was comparable to the other comparison conditions. OBS-COMP-HINT did not significantly differ in endorsement accuracy from OBS-SINGLE, OBS-COMP, and CLASS-COMP (all *p*s > .292); likewise, no simple interactions reached significance (all *p*s > .47). OBS-COMP-HINT led to increased endorsement accuracy relative to Baseline (β = 0.46, *SE* = 0.191, *Wald Z* = 2.411, *p* = .016). There was a significant effect of block in the OBS-COMP-HINT condition (β = 0.254, *SE* = 0.046, *Wald Z* = 5.533, *p* < .001), and a significant simple interaction where OBS-COMP-HINT led to smaller improvements across training relative to Baseline (β = -0.205, *SE* = 0.064, *Wald Z* = -3.207, *p* = .001). This suggests that providing a hint at the onset of training did not improve endorsement accuracy (relative to a comparable condition undergoing spontaneous transfer). It also provided further evidence that observational and classification comparison tasks produce comparable transfer benefits (*cf.* Patterson & Kurtz, 2020). Given that the spontaneous transfer effects occurred at such an early point of training, a hint may not be required to facilitate transfer in the comparison tasks.

A second hint-aided transfer analysis explored the effect of providing a hint in the midst of learning (i.e., at block 4) relative to continued first-time (Baseline) and hint-aided (OBS-COMP-HINT) learning (Figure 4). There were no significant differences between conditions (all *p*s > .145), however, we note a non-significant trend for greater accuracy in CLASS-COMP relative to OBS-SINGLE (β = 0.535, *SE* = 0.294, *Wald Z* = 1.818, *p* = .069). There was no effect of block for Baseline, OBS-SINGLE, CLASS-COMP, and OBS-COMP-HINT (all *ps* > .393), yet accuracy increased across training in the OBS-COMP condition (β = 0.141, *SE* = 0.055, *Wald Z* = 2.568, *p* = .01). Increased improvement across training in the OBS-COMP condition was reflected via a significant simple interaction term between Baseline and OBS-COMP (β = 0.179, *SE* = 0.075, *Wald Z* = 2.375 , *p* = .018), such that there were larger increases in accuracy across training in the OBS-COMP condition. We note similar non-significant trends for OBS-COMP relative to OBS-SINGLE (β = 0.126, *SE* = 0.074, *Wald Z* = 1.698, *p* = .09) and CLASS-COMP (β = 0.135, *SE* = 0.077, *Wald Z* = 1.759, *p* = .079). All other simple interaction terms were non-significant (all *p*s > .206). These results failed to demonstrate that a mid-training hint improved accuracy; except for the OBS-COMP condition where a hint may have enabled continued learning. The limited evidence of the hint's effectiveness in only the OBS-comp condition is loosely consistent with the cued-transfer advantage in the observational comparison condition observed by Patterson and Kurtz (2020).

Discussion

The results supported our main prediction that comparison-based encoding would improve spontaneous transfer of relational categories. We found evidence of spontaneous transfer in the early TCL blocks of both OBS-COMP and CLASS-COMP relative to Baseline. Along with the results of Snoddy and Kurtz (2020), this provides further evidence of the viability to study transfer through category learning tasks and supports extending our theoretical conception of transfer from a one-shot application of knowledge to solve a problem (e.g., Gick & Holyoak, 1980. 1983) to using prior knowledge to facilitate learning in a new domain.

In contrast to our predictions, OBS-SINGLE did not differ from Baseline or the comparison tasks. These results occurred despite category learning tasks attaining comparable levels of mastery during BCL, which suggests that the extent to which a category learning task supports schema abstraction may be a critical factor in spontaneous transfer success. When BCL and TCL tasks were congruent, Snoddy and Kurtz (2020) found evidence of transfer in a single-item presentation—a task which provides limited support for schema abstraction. In the present experiment, BCL and TCL phases involved different category learning tasks, which creates another dimension over which knowledge must be transferred (Barnett & Ceci, 2002). With the lack of congruent BCL and TCL tasks, the limited abstraction afforded by single-item presentations may not be sufficient to effectively support transfer. Increased support for schema abstraction in the comparison tasks may have better facilitated access and led to only those tasks showing evidence of transfer relative to Baseline.

Such an interpretation builds upon the cued-transfer effects of Kurtz et al. (2013) and Patterson and Kurtz (2020) to demonstrate that comparison-based category learning tasks can also promote spontaneous transfer. Further, these results suggest that comparison plays a similar role in analogical transfer (e.g., Gick & Holyoak, 1983) and the transfer of knowledge to facilitate performance in a subsequent category learning task—comparison supports the spontaneous access of prior knowledge and subsequent transfer. Future work should explore whether BCL and TCL task congruence mediates transfer effects in single-item presentations.

The present findings appear inconsistent with Patterson and Kurtz (2020), as observational and classification comparison tasks produced comparable performance on both transfer tasks (although there was a marginal effect during

hint-aided transfer). However, Patterson and Kurtz (2020) found that observational comparison led to improved old-item test performance, which suggests that classification-based comparison may result in slower acquisition that may have been detrimental to performance on subsequent tasks. The present experiments increased the length of BCL training to avoid such differences in initial learning, which may have led to comparable performance between observational and classification comparison during the transfer assessment. As the present experiment involved a spontaneous transfer assessment and training-based assessments (as opposed to unsupervised test trials), the present results may better reflect a boundary condition on the benefits of observational comparison performance as opposed to a failure to replicate Patterson and Kurtz (2020).

In contrast to analogical transfer experiments (Gick & Holyoak, 1980, 1983), we failed to find robust evidence of a hint improving transfer both when a hint was provided at the onset of learning and when it was provided in the midst of TCL training (aside from limited evidence that a hint allowed for OBS-COMP to continue learning). As the present study strengthened the hint relative to Snoddy and Kurtz (2020), it further demonstrates an incompatibility between hint-aided transfer and the category learning approach (for more on this incompatibility, see Snoddy and Kurtz, 2020). As the spontaneous transfer effects occurred early in TCL training, the structural similarity between domains may have been readily apparent and obviated the benefit of a hint. This suggests that spontaneous access may be easier to achieve when transferring knowledge acquired during one category learning task to another than during problem-solving. However, we cannot rule out that the ease in access stems from the particular category structures and stimulus sets used in the present experiment. Future work should explore whether similar effects emerge when using other categories.

The recognition data provided partial support for our hypotheses—all category learning conditions led to a significantly higher rate of false alarms than the Baseline condition, yet there were no differences amongst the category learning conditions. Such an effect is consistent with the results of Snoddy and Kurtz (2020)—a prior category learning opportunity, regardless of the task, led to increased false alarms (and had no consistent effect on hit rate). As in Snoddy and Kurtz (2020), this reflects a partial extension of the results of De Brigard et al. (2017) to relational categories.

The different pattern of results between TCL and recognition phases (i.e., no effect of comparison on recognition memory) suggest that schema abstraction does not play a critical role in false alarms. This suggests that learning about the category structures, regardless of whether an abstract representation is formed, may be sufficient to increase false alarms through familiarity (Yonelinas, 2002). Such a familiarity-based pathway to recognition may result from participants experiencing a structure-based reminding of BCL items or categories (Ross, 1984) or through categorizing the BCL items with respect to known categories. An alternative explanation of this effect is that BCL may serve to prime known relational structures and increase the rate of false alarms (Popov, Hristova, & Anders, 2017). Future work should further explore the mechanism behind such recognition effects in relational category learning.

References

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128,* 612 – 637.

De Brigard, F. D., Brady, T. F., Ruzic, L., & Schacter, D. (2017). Tracking the emergence of memories: A category-learning paradigm to explore schema-driven recognition. *Memory and Cognition*, *45,* 105-120.

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., 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.

Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1303-1310.

Markman, A. B., & Gentner, D. (2000). Structure-mapping in the comparison process. *American Journal of Psychology, 113,* 501-538.

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science, 9,* 247-255.

Patterson, J. D., & Kurtz, K. J. (2020). Comparison-based learning of relational categories (you'll never guess). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46,* 851-871.

Popov, V., Hristova, P., & Anders, R. (2017). The relational luring effect: Retrieval of relational information during associative recognition. *Journal of Experimental Psychology: General, 146,* 722-745.

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16,* 371-416.

Snoddy, S., & Kurtz, K. J. (2020). Analogical transfer and recognition memory in relational classification learning. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 81-87). Cognitive Science Society.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441-517.