Analogical Transfer and Recognition Memory in Relational Classification Learning

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

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

Abstract

People spontaneously make connections between superficially distinct domains through relational similarity, but this spontaneous transfer has yet to be demonstrated across distinct classification tasks. A related issue is that the acquisition of a category may affect recognition memory for category-consistent items. Participants in the Category Learning condition completed an initial classification task. The Category Learning and Baseline conditions each received category-consistent items to study followed by a recognition test. Both groups completed a final classification task in a novel domain abiding by the same underlying category structures as the initial classification task. The Category Learning group showed 1) increased false alarms during the recognition test and 2) higher accuracy in the final classification task (when told the classification phases were unrelated). This suggests that classification learning led to a schematization of the category-defining concept (evidenced by increased false alarms), which supported spontaneous transfer of relational concepts across distinct classification tasks.

Keywords: analogical transfer; relational categories; classification; recognition memory

Introduction

Relational concepts are adaptive because they can be instantiated within multiple domains. Such analogical transfer and its limitations was demonstrated in a seminal study by Gick and Holyoak (1980). Participants first read a base passage about a general who captured a fortress by dividing an army into many small groups that simultaneously attacked the fortress from multiple angles. Participants were then given a target problem, which described a doctor who had a patient with an inoperable tumor and a ray device that could destroy the tumor but would harm the healthy tissue. Participants were asked to devise a solution to destroy the tumor without harming the patient. The critical question was whether the relational concept (convergence) demonstrated by the base passage would be transferred to the target problem to create a solution (i.e., firing less intense rays from around the tumor simultaneously). Participants who received a hint about the relevance of the base passage readily transferred the solution. When the base passage and target problem were presented as separate experiments, which requires noticing the analogy between them, few participants were able to successfully transfer the concept. This pattern of results highlights the key impediment to successful transfer—the frequent failure to spontaneously access germane, but superficially dissimilar prior knowledge.

As an analogy and a relational category are both structured by a relational concept (i.e., all members of a relational category are analogous to one another) (Gentner & Kurtz, 2005), recent work has leveraged the categorization literature to design novel study tasks that support spontaneously accessing analogically similar prior knowledge. Kurtz and Honke (2019) demonstrated that category construction is an effective way to encourage spontaneous transfer during problem-solving tasks. Likewise, Snoddy and Kurtz (2017) demonstrated that category-building, an enhanced form of sequential summarization of category relevant information, is effective in increasing the accessibility of analogically related prior knowledge in memory. Given the utility of categorization-inspired study tasks in promoting analogical transfer during problem solving tasks, we ask whether categorization may be fruitful in another way: assessing spontaneous transfer as using one category learning task to enhance learning of the same underlying relational category structure in another domain.

Cued transfer has already received some attention in the categorization literature. Casale, Roeder, and Ashby (2012) operationalized transfer as an extension of a learned category structure to an unexperienced region of feature space during classification training. Casale et al. (2012) demonstrated successful application of a rule-based category to novel items without a loss of classification accuracy, however, participants had to relearn an information integration category structure for novel items—a failure to extrapolate that category structure to a new region of the feature space. This study used attribute-based categories, which are defined by members' featural similarities (Gentner & Kurtz, 2005), so both the training and transfer items are from the same domain (albeit with different ranges of feature values). Such within-domain extrapolation is much less demanding than analogical transfer as it does not involve spontaneously accessing relevant prior knowledge from a superficially dissimilar domain (Gentner, Rattermann, & Forbus, 1993).

Cued analogical transfer between relational category learning tasks has been explored in a study by Kurtz, Boukrina, and Gentner (2013). Participants were assigned to either comparison-based or single-item classification tasks to acquire three relational categories (see Figure 1)—support, monotonicity, and symmetry—through supervised feedback. Participants underwent classification training followed by an unsupervised test including within-domain generalization to novel items. An analogical transfer assessment presented items in a new domain (mobile-like displays) that abided by the same three relational categories and asked participants to classify each display according to the same labels used during training. Comparison led to improved transfer relative to single-item; both conditions provided evidence of analogical transfer with above-chance performance (Kurtz et al., 2013). Both Casale et al. (2012) and Kurtz et al. (2013) used the same category labels and contextual continuity (all phases of the same experiment), which attenuates the demands of accessing an analogous source.

Prior work has demonstrated that comparison of multiple analogs at the time of study promotes spontaneous transfer in the problem solving task by encouraging the formation of an abstract representation of the shared relational concept that is more accessible in memory than representations of individual analogs (Gick & Holyoak, 1983). Recent work by De Brigard, Brady, Ruzic, and Schacter (2017) tested a related hypothesis—category learning leads to schematization of the category-defining concept (i.e., an abstract representation), which affects recognition memory for category-consistent items. De Brigard et al. (2017) demonstrated increases in both the hit rate for old items and the false alarm rate for lures that were consistent with a learned category based on a unidimensional rule. As De Brigard et al. (2017) used an attribute-based category, this effect could emerge from either schematization of the category structure during learning or learning to attend to a perceptually identifiable feature.

In the present study, we sought to investigate a classification task suitable for assessing spontaneous analogical transfer. As in past studies, an initial classification task provided an opportunity to acquire a set of relational categories and a final classification task was used to measure transfer. The final classification task was presented under the guise of a new experiment (Gick & Holyoak, 1980), the category structures were instantiated in a novel domain (Kurtz et al., 2013), and new category labels were used. For spontaneous transfer to occur in this task, participants must: 1) access knowledge of the category structures from initial classification learning; 2) evaluate the structures in both tasks as analogous; and 3) apply the retrieved knowledge to promote successful classification learning in the new domain. In this approach, transfer is evidenced by improved accuracy during the final classification task relative to experiencing the same task without prior learning to draw upon.

A second goal of the present study is to test whether the effect of category learning on recognition memory extends to relational categories, which are often characterized as 'rule-like' (Gentner & Kurtz, 2005) as membership is based on whether a given stimulus conforms to the category-defining system of relationships. If the findings of De Brigard et al. (2017) are due to schematization of the category structure, then they will extend to relational categories, which are not based on the perceptually identifiable features of a stimulus. By including a recognition memory phase in the transfer classification task, future work can explore whether such schematization effects on memory are correlated with spontaneous transfer (see Discussion for an explanation of why this was not appropriate in the present study).

This two-part classification measure reflects a novel way to examine spontaneous analogical transfer. This measure is related to the standard approach of using problem solving assessments (e.g., Gick & Holyoak, 1980)—both involve forming a representation of a concept through a study task then spontaneously accessing that representation and applying it to a new domain during the transfer assessment—yet offers a broader conception of analogical transfer. The standard approach has focused on acquisition of a single concept through exposure to a few cases paired with a high-impact task such as comparison (Gick & Holyoak, 1983). The classification-based approach extends acquisition to multiple concepts being jointly considered with a larger number of study items and an extended period of learning. Classification also differs in that it is an iterative learning task, which involves making a guess about category-relevant information for each item then updating the representation of the category based on corrective feedback (Patterson & Kurtz, 2019).

The way analogical transfer occurs may differ between approaches. The standard approach views transfer as a single process—it involves noticing that the problem is analogous between the representation of the study cases and the transfer problem, then applying the solution from the representation of the study cases to the transfer problem through inference (Gentner, 2010). Transfer may be similar in the classification approach—it could involve immediately noticing that the relational concepts are analogous between domains and directly applying the concepts to make sense of the transfer domain. Alternatively, it may be an iterative process that requires learning how to transfer (Bransford & Schwartz, 1999). This may involve learning how analogous relations are instantiated within the transfer domain or how to align the instantiation of a relation in one domain into another.

The classification approach is robust to differences in how concepts are represented. The standard approach often uses textual stimuli along with acquisition and transfer tasks that encourage a verbal description of the concept, which focuses on the transfer of explicit representations and has largely overlooked the transfer of implicitly represented concepts (Day & Goldstone, 2011). It has not yet been established how relational category structures like those used in Kurtz et al. (2013) are represented, but they appear compatible with both explicit representations (e.g., "support") and implicit representations (e.g., the common spatial arrangement of elements within each category). Casale et al. (2012) have demonstrated that classification tasks can support acquisition of both explicit and implicit representations of concepts, so the classification approach is amenable to evaluating transfer of both implicit and explicit representations of concepts.

Present Study

The present study had two distinct, complementary goals: 1) to provide evidence of spontaneous analogical transfer in a classification task, and 2) to test whether the acquisition of a relational category affects recognition memory through schematization of knowledge. The procedure involved two classification phases: 1) an initial classification task that provided an opportunity to acquire the categories and 2) a final classification task that was introduced as a new experiment (with different category labels and stimulus domains) to measure spontaneous transfer. A recognition memory phase was included between the classification tasks. Participants received novel, category-consistent items to study, then took a recognition test containing items from the study set intermixed with novel, category-consistent lures.

Participants assigned to the Category Learning condition took part in all phases of the experiment. During final classification, the Category Learning group was randomly assigned to either a Spontaneous condition, which did not indicate any relationship between the two classification phases, or a Hint-aided condition which specified that the two classification tasks were related and abided by the same category structures. The Baseline condition did not receive initial classification, so the performance of this group reflects the absence of category learning for both the recognition memory and final classification phases.

We predicted that both Spontaneous and Hint-aided conditions would demonstrate analogical transfer in the final classification learning task. This could be realized by improved accuracy during the final classification task in the Spontaneous and Hint-aided conditions relative to either the Baseline condition or a within-subject baseline of initial classification performance. Transfer differences between Spontaneous and Hint-aided conditions were exploratory. To the extent that the classification procedure invokes similar processes and representations as the problem-solving task, then the provision of a hint should mitigate the difficulty of spontaneously accessing prior knowledge resulting in either higher accuracy overall or higher accuracy in early blocks of training in the Hint-aided condition than in the Spontaneous condition. If classification encourages schematization of knowledge, the Category Learning condition is predicted to result in a higher rate of hits and false alarms relative to the Baseline condition on the recognition memory test.

Method

Participants

A total of 167 undergraduate students from Binghamton University participated. Participants were randomly assigned to the Category Learning (N = 88) or Baseline (N = 79) conditions. During final classification, the Category Learning condition was randomly subdivided into Spontaneous (N = 46) and Hint-aided (N = 42) conditions.

Materials and Design

The stimuli consisted of 72 images from two distinct domains—rock arrangements and mobiles (see sample items in Figure 1). All 36 of the rock arrangements and 15 of the mobiles were from Patterson and Kurtz (2019; see also Kurtz et al., 2013). An additional 21 mobile stimuli were created so that the domains could be counterbalanced across phases. Stimuli varied in the superficial attributes of their individual elements (rocks or geometric shapes): size, shape, and color. None of the superficial attributes were predictive of category membership. Each of the stimuli demonstrated exactly one of the three possible category structures (12 stimuli per category and domain): 1) *support*—a rock/shape being supported by two other rocks/shapes, 2) *monotonicity*—a decrease in height from left to right, or 3) *symmetry*—two rocks/shapes that are stacked atop each other and similar in size, shape, and color. The same category structures were instantiated in each domain, but the Mobile domain was inverted over the x-axis and associated with different category labels (Zibble, Wuggy, and Doppa) than the Rock domain (Besod, Makif, and Tolar).

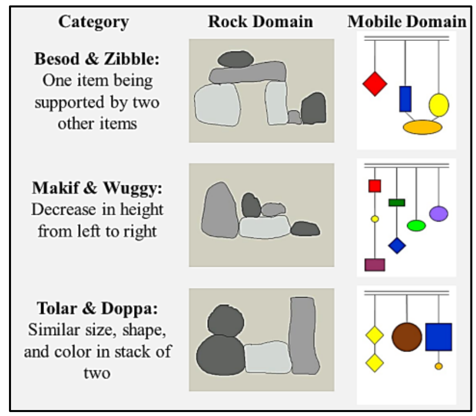

Figure 1: Sample stimuli for each category within the Rock and Mobile domains.

The domain presented in each phase was counterbalanced across participants. The Rock/Mobile domain was assigned to the initial classification and recognition memory phases while the Mobile/Rock domain was assigned to the final classification phase. For initial classification and recognition memory phases, 18 (6 per category) stimuli were randomly assigned to classification, 9 (3 per category) to the study component of the recognition memory phase, and 9 (3 per category) as new items were used as lures on the recognition test. For the final classification phase, 18 (6 per category) stimuli were randomly selected for each participant to equate for stimulus exposure in initial classification.

Procedure

Upon entering the lab, participants were told that they would be taking part in multiple experiments throughout the session. A research assistant then executed a PsychoPy (Peirce, 2007; version 1.83) program that presented the initial classification and recognition memory phases. Participants in the Category Learning condition first proceeded to the initial classification phase; the Baseline condition proceeded directly to the recognition memory phase. Participants in the Category Learning condition were presented with instructions that included a cover story, provided with the three category labels, told that the task involved learning what makes a given item belong to a category, and told that they will be tested on their knowledge of the categories later in the experiment. These instructions did not specify the nature of the test phase so that participants were not aware of the upcoming recognition memory phase during initial classification. Participants then underwent 5 blocks of classification for 18 items (6 per category). Within each block, participants were sequentially presented with each of the items in a randomized order. The three category labels were presented below each item and participants made a classification response by clicking on a category label. After making each classification response, participants received corrective feedback.

After the initial classification phase, participants proceeded to the recognition memory phase, which consisted of study and test sub-phases. Participants were told they were now in the "study phase" of the experiment and instructed to study each item so that they would be able to recognize whether it belonged to this "study set" later. Nine new items (not included in initial classification) were sequentially presented to participants without category labels. Once an item appeared on the screen, participants were able to study it for as long as they wished before proceeding to the next item.

Immediately after the study phase, the recognition test was administered. Participants were told that they would be shown a series of items and their task was to decide whether the item was from the "study phase". Participants were sequentially presented with 18 items—9 items from the study phase intermixed with 9 new lures (not presented during study or initial classification phases)—and made a yes/no judgment about whether each item was from the study phase. No feedback was given after each response. After the recognition assessment, the PsychoPy program was closed by the research assistant and participants were told that they would be moving on to the next experiment.

All participants then proceeded to final classification. To make the classification tasks appear unrelated, this phase was introduced as a new experiment by the research assistant and a different PsychoPy program was executed to present the final classification task. To further enhance this context shift, superficial changes were made to the final classification experience: 1) participants received a generic version of the instructions from initial classification without the cover story; 2) the category structures were instantiated in a different domain; 3) different category labels were used; and 4) the background colors of the stimuli, experiment program, and response buttons were all changed between classification phases. Spontaneous and Baseline conditions received instructions stating that they "will be shown images that belong to one of three categories" and their "goal was to learn how to correctly identify which category each image belongs to". The Hint-aided condition received an additional instruction stating that "the categories from the previous experiment (initial classification phase) would be helpful in learning the present categories". Participants in all conditions then underwent 5 blocks of supervised classification for 18 items following the same procedure as initial classification.

Results

Initial Classification Learning

No predictions were made regarding initial classification, but differences in this phase would bar comparisons between experimental conditions and certain baselines. The first question was whether the recognition memory experiment affected subsequent classification performance. To address this question, performance was compared between Category Learning (initial classification) and Baseline (final classification) conditions using a mixed-effects logistic regression model that allowed subject and item to vary as random intercepts and predicted trial-wise accuracy with condition, block, and their interaction. There was a significant effect of block for both the Baseline (β = 0.321, SE = 0.02, Wald Z = 16.168, p < .001) and Category Learning (β = 0.364, SE = 0.019, Wald Z = 18.858, p < .001) conditions, such that accuracy increased as training progressed. There was no significant difference between conditions (β = 0.034, SE = 0.172, Wald Z = 0.189, p = .843) and no significant interaction between condition and block (β = 0.043, SE = 0.028, Wald Z = 1.552, p = .121) (Figure 2). This suggests that the inclusion of a recognition memory phase did not significantly alter performance in a subsequent classification task and that the Baseline condition serves as a valid comparison in final classification analyses.

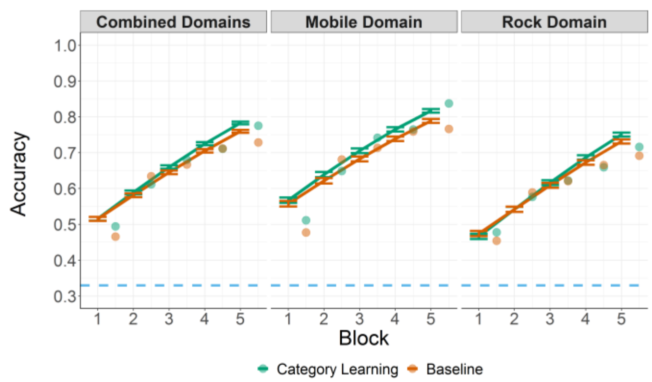

Figure 2: Each line reflects adjusted mean classification accuracy (right-offset points are unadjusted means) across blocks of training. Error bars are ± 1 SEM and the dashed line reflects chance performance.

The second question was whether the two stimulus domains resulted in comparable classification accuracy when first acquired. A mixed-effects logistic regression model that allowed subject and item to vary as random intercepts and predicted trial-wise accuracy with domain, block, and their interaction was used to address this question. There was a significant effect of block for both the Rock (β = 0.281, SE = 0.019, Wald Z = 15.129, p < .001) and Mobile (β = 0.42, SE = 0.021, Wald Z = 20.067, p < .001) domains, such that accuracy increased as training progressed. While there was no significant difference in accuracy between domains (β = 0.089, SE = 0.196, Wald Z = 0.452, p = .651), there was a significant interaction between block and domain (β = 0.14, SE = 0.028, Wald Z = 5.001, p < .001) such that the Mobile domain was associated with greater increases in accuracy across blocks than the Rock domain. This confound between domain and accuracy bars the comparison of final classification performance and the within-subject baseline (initial classification in the Category Learning condition).

Recognition Memory

The prediction for the recognition test was that the Category Learning condition would encourage schematization of knowledge thereby increasing both the rate of hits for items contained in the study phase and the rate of false alarms to category-consistent lures relative to the Baseline condition. A mixed-effects logistic regression model that allowed subject and item to vary as random intercepts and predicted trial-wise accuracy with condition was used to test this prediction. In contrast to the prediction, there was no significant difference in the proportion of hits between the Category Learning and Baseline conditions ($\beta = 0.146$, $SE = 0.188$, Wald $Z = 0.775$, $p = .438$). In support of the prediction, the Category Learning condition led to a significantly higher rate of false alarms than the Baseline condition ($\beta = 0.709$, $SE = 0.217$, Wald $Z = 3.362$, $p = .001$) (Figure 3).

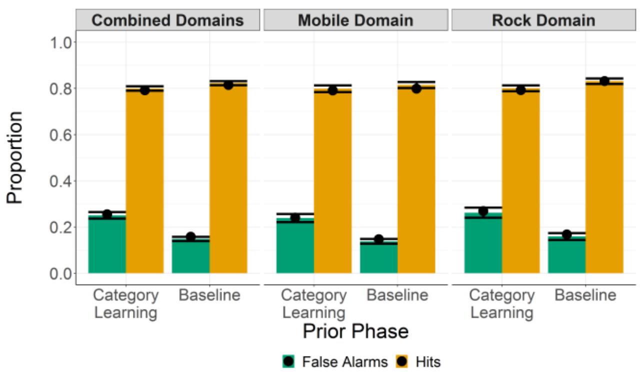

Figure 3: Bars reflect adjusted mean proportion of hits and false alarms (points are unadjusted means) as a function of condition. Error bars are ± 1 SEM.

Supplemental analyses added domain and the interaction between condition and domain into the regression models to test whether differences in the ease of acquisition between domains affected recognition memory performance. The lack of significant differences in hits between Category Learning and Baseline conditions and increased rate of false alarms in the Category Learning condition both persisted after controlling for domain. With respect to hits, there was no significant difference between stimulus domains ($\beta = 0.0002$, $SE = 0.257$, Wald $Z = 0.001$, $p = .999$) or interaction between domain and condition ($\beta = 0.25$, $SE = 0.376$, Wald $Z = 0.666$, $p = .505$). Like-wise, for false alarms there was no significant difference between domains ($\beta = 0.15$, $SE = 0.345$, Wald $Z = 0.434$, $p = .664$) or interaction between domain and condition ($\beta = 0.062$, $SE = 0.433$, Wald $Z = 0.143$, $p = .886$).

Given the difference in the rate of false alarms, a remaining question was whether the Category Learning led to a response bias. To address this concern, each participant's sensitivity and bias were calculated. As these measures aggregate across trial-wise responses, a linear regression model with condition as a predictor was used in lieu of a mixed-effects approach. Overall, Category Learning ($M = 1.409$, $SD = 0.869$) led to significantly lower sensitivity than Baseline ($M = 1.742$, $SD = .692$) ($\beta = 0.333$, $SE = 0.122$, $t = 2.718$, $p = .007$), which is driven by the increased rate of false alarms in the Category Learning condition. With respect to bias, Category Learning ($M = 1.336$, $SD = .914$) did not significantly differ from Baseline ($M = 1.588$, $SD = 1.059$) ($\beta = 0.252$, $SE = 0.153$, $t = 1.653$, $p = .1$). Overall, this suggests that forming a representation of a category led to an increase in falsely recognizing category-consistent lures, and that this effect is not attributable to a response bias.

Final (Transfer) Classification Performance

The prediction for the final classification task was that the Category Learning conditions (Spontaneous and Hint-aided) would have higher classification accuracy than participants learning the category structures for the first time (Baseline). Either differences between conditions or an interaction between condition and block would provide evidence of analogical transfer. A mixed-effects logistic regression model that allowed both subject and item to vary as random intercepts and predicted trial-wise accuracy with condition, block, and their interaction was used to address this question. There were no significant differences in accuracy between conditions: Spontaneous and Baseline ($\beta = 0.002$, $SE = 0.205$, Wald $Z = 0.01$, $p = .992$), Hint-aided and Baseline ($\beta = 0.281$, $SE = 0.21$, Wald $Z = 1.334$, $p = .182$), and Spontaneous and Hint-aided ($\beta = 0.278$, $SE = 0.236$, Wald $Z = 1.178$, $p = .238$). There was a significant effect of block for Spontaneous ($\beta = 0.444$, $SE = 0.028$, Wald $Z = 15.763$, $p < .001$), Hint-aided ($\beta = 0.284$, $SE = 0.028$, Wald $Z = 10.246$, $p < .001$), and Baseline ($\beta = 0.318$, $SE = 0.02$, Wald $Z = 16.087$, $p < .001$) conditions. In support of the prediction regarding transfer, there were significant interactions between block and the Spontaneous and Baseline ($\beta = 0.126$, $SE = 0.034$, Wald $Z = 3.671$, $p < .001$) conditions and between block and the Spontaneous and Hint-aided conditions ($\beta = 0.16$, $SE = 0.039$, Wald $Z = 4.044$, $p < .001$), such that there were larger improvements in the Spontaneous condition across block than in the Baseline and Hint-aided conditions. There was no significant interaction between block and the Hint-aided and Baseline conditions ($\beta = -0.034$, $SE = 0.034$, Wald $Z = -0.988$, $p = .323$) (Figure 4).

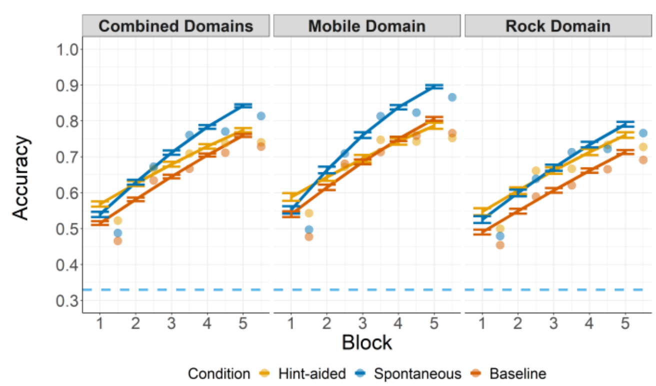

Figure 4: Each line reflects adjusted mean classification accuracy (right-offset points are unadjusted means) across blocks of training. Error bars are ± 1 SEM and the dashed line reflects chance performance.

Supplemental analyses applied the previous model to each domain individually (as the model including domain as another predictor failed to converge) to test whether the observed transfer effects were consistent across domains. For the Mobile domain, the lack of differences between conditions (all *p*s > .3) and the significant effect of block were replicated (all *p*s < .001). Within the Mobile domain, the significant interactions between block and the Spontaneous and Baseline (β = 0.186, *SE* = 0.053, Wald *Z* = 3.515, *p* < .001) conditions, the significant interaction between block and the Spontaneous and Hint-aided conditions (β = 0.269, *SE* = 0.059, Wald *Z* = 4.540, *p* < .001), and the lack of a significant simple interaction between block and the Hint-aided conditions and Baseline condition (β = 0.082, *SE* = 0.049, Wald *Z* = 1.672, *p* = .094) were all replicated. For the Rock domain, the lack of differences between conditions (all *p*s > .381) and the significant effect of block was replicated (all *p*s < .001). None of the simple interactions were significant, but we note a non-significant trend where the Spontaneous condition led to greater improvement across blocks than the Baseline condition (β = 0.087, *SE* = 0.045, Wald *Z* = 1.908, *p* = .057; all other *p*s > .134). The improvements in later blocks of final classification performance for the Spontaneous condition appear to be driven by the Mobile domain.

Discussion

The aim of the present study was to test the following predictions: 1) acquisition of a relational category would increase the rate of both hits and false alarms in a recognition memory task, and 2) category learning would facilitate performance in the final classification task (i.e., analogical transfer). Results from the recognition memory assessment provided partial support for the predictions. The Category Learning condition led to an increased rate of false alarms to category-consistent lures when compared to the Baseline condition, but the Category Learning condition did not affect the rate of hits for old items. Results for the final classification phase provided partial support for the prediction regarding analogical transfer. Final classification accuracy was higher in the Spontaneous condition during later blocks of classification relative to Hint-aided and Baseline conditions. While this effect was primarily driven by the Mobile domain (we note a marginal trend relative to Baseline in the Rock domain), it suggests that participants were spontaneously accessing knowledge of analogically similar categories from initial classification and transferring it to the final classification task to facilitate performance.

The finding that category learning increased false alarms to category-consistent lures suggests that acquisition of a relational category encourages schematization of the underlying structure, which may facilitate attention to the category-defining structure (De Brigard et al., 2017). An increased attentional focus when processing both new and old items could support the detection of a structural match between these items and known categories, which could create a sense of familiarity, increasing hits and false alarms (Yonelinas, 2002). The absence of category learning in the Baseline condition could result in the default bias to focus on the superficial features of the stimuli (Kotovsky & Gentner, 1996). Such a superficial bias would support hits through recollection based on superficial details of stimuli, but not false alarms as each item had unique, albeit similar, superficial details (Yonelinas, 2002). Alternatively, the lack of differences in hit rate may be attributable to a ceiling effect. Future work should replicate this finding with more items in the recognition memory phase and a brief distractor task to determine if a difference in hits can be observed.

The increase in false alarms arising from familiarity-based recognition is related to the relational luring effect (RLE)—priming known semantic relations during study increases the amount of time taken to correctly reject novel, analogous items in an associative recognition task (Popov, Hristova, & Anders, 2017). The RLE assumes an abstract representation of relations in memory (i.e., schematized relations) and that priming these relations during study can create a sense of familiarity to novel, relationally similar items at test (Popov et al., 2017). If initial classification only serves to activate known relations, the RLE suggests that a priming mechanism is driving the increased rate of false alarms. Alternatively, if the concepts were acquired during initial classification or participants learned how to adapt known concepts into specific domains, familiarity-based recognition could also be achieved through either the perception of structural similarity between the known category and recognition item or classification of the recognition item as being a member of a recently acquired category. Future work should seek to disambiguate whether priming, similarity-based retrieval, or categorization are responsible for increased false alarms.

The evidence of transfer in only the Spontaneous condition appears counterintuitive as the provision of a hint in problem-solving tasks generally facilitates transfer by attenuating the access demands so that transfer primarily involves application of the relational concept (Gick & Holyoak, 1980). A key difference between problem solving and classification-based measures of transfer is that the classification approach allows for participants to learn how to instantiate a concept within a novel domain. As the spontaneous transfer effect emerged during later blocks of classification, it is plausible that some initial exploration or learning of the categories in the novel domain was required for transfer. Such experience could support induction of the category structure, which facilitates retrieval of categorical knowledge. It could also provide domain experience, so that when relevant knowledge is retrieved it can be readily applied. As the Hint-aided condition was told that the known categories would be helpful in learning this domain, they may have only focused on applying those categories to a new domain at the expense of initial exploration that may have provided additional insights about the category structures. In contrast to the application of a solution to a new domain, understanding how to instantiate a category structure within a new domain may be comparably difficult as acquiring it for the first time, which would result in the lack of evidence for transfer in the

Hint-aided condition. Future work should test whether this is a meaningful difference between transfer measures or a mistranslation of the hint manipulation by using a stronger hint that provides direct instruction of the category structures and how to apply them to a new domain.

Alternatively, the lack of transfer in the Hint-aided condition may suggest representational differences between the present category structures and the kinds of principles often used in the standard approach. Day and Goldstone (2011) demonstrated that analogical transfer of an implicit representation between two domains does not benefit from the provision of a hint and that participants' awareness of the explicit solution strategy could not fully account for successful transfer. If the relational categories in the present study were represented implicitly, the provision of an explicit hint may have impaired the access and/or application of relevant prior knowledge. Future work should explore the types of representations formed when learning the relational categories used in the present study and how different representations may mediate the effectiveness of a hint.

A limitation of the present work is that the Mobile domain was easier to learn than the Rock domain. The difference in ease of acquisition between domains may have been driven by the Rock domain having a higher degree of superficial similarity between the elements of each arrangement (e.g., a large circular dark rock vs. a small elliptical light rock) than the Mobile domain (e.g., a yellow triangle vs. a red square) or that the 'wires' included in the Mobile domain may have provided a stronger cue to the system of relations reflecting the category structures than the Rock domain. Initial learning of the more difficult Rock domain may have led to better transfer to the easier Mobile domain, however, initial learning in the easier Mobile domain may not have led to a robust enough understanding of the category structures to support transfer to the more difficult Rock domain.

This confound between ease of acquisition and domain prohibited the use of a within-subject measurement of transfer (final – initial classification) in the present study. The lack of a within-subject transfer metric prevented testing a third prediction— that false alarm rate would be predictive of transfer (i.e., analogical transfer and false recognition may be driven by the same mechanism). When evaluating final classification accuracy in the Category Learning condition, any observed effect could not be clearly attributable to either transfer or another aspect of classification performance. Future work should create additional domains for the present category structures that are normed for difficulty so that relationships between memory and spontaneous transfer can be assessed.

References

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24,* 61-100.

Casale, M. B., Roeder, J. L., & Ashby, F. G. (2012). Analogical transfer in perceptual categorization. *Memory and Cognition, 40,* 434-449.

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

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

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In Ahn, W. L., Goldstone, R. L., Love B. C., & Wolff, P. W. (Eds.), *Categorization inside and outside the lab.* 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.

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

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.

Kurtz, K.J., & Honke, G. (2019). Sorting out the problem of inert knowledge: Category construction to promote spontaneous transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication. https://doi.org/10.1037/xlm0000750

Patterson, J. D., & Kurtz, K. J. (2019). Comparison-based learning of relational categories (you'll never guess). *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0000758

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods, 162,* 8-13.

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.

Snoddy, S., & Kurtz, K. J. (2017). Promoting spontaneous analogical Transfer: The role of category status. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 2832-2838). Austin, TX: 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.