Simulating feature- and relation-based categorisation with a symbolic-connectionist model

Simulating Feature- and Relation-Based Categorisation with a Symbolic-Connectionist Model

Ekaterina Y. Shurkova (e.shurkova@ed.ac.uk)
Leonidas A. A. Doumas (alex.doumas@ed.ac.uk)
Department of Psychology, University of Edinburgh
7 George Square, Edinburgh EH8 9JZ, Scotland, UK

Abstract

Participants in Goldwater et al. (2018) reported using either feature- or relation-based strategy during a series of category learning tasks. A computational modeling study was conducted to investigate whether performance on Experiments 1 and 2 of Goldwater et al. (2018) might be explained by the assumption that participants used either feature- or relation-based representational encoding during learning. Human participants' and model performance are compared and implications are discussed.

Keywords: categorisation; relational categories; featural categories; symbolic-connectionist model; computational modeling

There is an ongoing debate on whether featural and relational categorisation rely on the same cognitive mechanisms. Membership in a featural category is assessed via similarities of objects' intrinsic features, while membership in a relational category is evaluated based on relations between parts of the stimulus. Some findings suggest that feature-based categorisation is fundamentally different from that based on relations (e.g., Davis, Goldwater, & Giron, 2017; Jung & Hummel, 2015), while other results demonstrate similarities in the two kinds of categorisation (e.g., Corral, Kurtz, & Jones. 2018; Higgins & Ross, 2011).

Goldwater et al. (2018) aimed to investigate this controversy by exploring different strategies that human participants might employ during category learning. The findings of their Experiments 1 and 2 suggest that during categorisation tasks participants differed on type of learning strategy. Some participants utilised a feature-based strategy while others applied a relation-based strategy.

One of the limitations of the study is the fact that the learning strategy is assessed with self-reports. This approach involves a higher-level metacognitive assessment of the strategies involved in task completion; an approach which may not properly identify the underlying cognitive mechanisms. We sought to further explore this issue using a computational model of relational reasoning (DORA; Doumas, Hummel, & Sandhofer, 2008).

The goal of the current study was to investigate whether feature- and relation-based strategies might be the result of engaging different representational encodings—by encoding each stimulus purely with its features or with a relational structure. In addition, the current study aimed to explore whether between- and within-category comparisons affect the model's performance on the categorisation task differently while either feature- or relation-based encodings are engaged. DORA (as its predecessor LISA) is one of a few models that can simultaneously represent featural and symbolic relational representations within the same framework. This is the main reason why this framework was employed to explore the difference between featural and relational categorisation.

Method

Category Learning Task

Stimuli for the simulations were adapted from Goldwater et al. (2018) and comprised two artificial categories. Each stimulus contained three lines formed by coloured squares—red, green, blue, and yellow. The distribution of colours served as the featural aspect of the stimulus, with more reds and greens probabilistically defining membership of one category, Snarg, and more blues and yellows defining another category, Blicket. The relational aspect of the stimuli, the relative line lengths, was deterministic of the category membership such that monotonically decreasing or increasing line lengths defined snargs and non-monotonic relations represented blickets (see Figure 1).

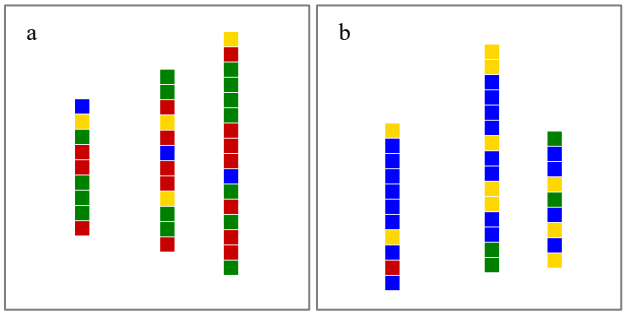

Figure 1: a) An example of a snarg; b) An example of a blicket; adopted from Goldwater et al. (2018).

The categorisation task consisted of a learning phase with feedback and a testing phase without feedback. During the learning phase the stimuli from different categories were presented one at a time. Upon classification of each stimulus the correct/incorrect feedback was given.

The testing phase comprised baseline, feature, relation, and cross-mapped trials. During the baseline test trial, the stimuli

were presented in the same manner as in the learning phase. For the feature test trial only featural aspects of the stimuli, colour distributions, were preserved, while the relational information, relative line lengths, was removed (by generating lines of equal length). During the relation test trial the colours were equally distributed, making category membership based on features impossible to discern, while the relative line lengths were preserved. For the cross-mapped test trial the colour distributions referred to one category while the relative line lengths referred to the other (see Goldwater et al., 2018, for full details of the stimuli and the learning and testing phases).

A Symbolic-Connectionist Model of Relational Reasoning

To explore the findings of Goldwater et al. (2018) we used DORA (Discovery Of Relations by Analogy; Doumas, Hummel, & Sandhofer, 2008), a symbolic-connectionist model previously shown to account for various learning and representation development phenomena. As input DORA takes objects represented in a distributed fashion as a collection of features and learns single-place predicates and multi-place relations (Doumas et al., 2008; Doumas & Martin, 2019). Ultimately, the model learns a relational schema which does not rely on features and may be used for analogical reasoning and generalisation (Doumas & Hummel, 2005; Doumas & Hummel, 2010; Hummel & Holyoak, 1997; Hummel & Holyoak, 2003).

DORA consists of an active memory (AM), the stimuli in the focus of DORA's "attention", and long-term memory storage (LTM), where all the information "seen" previously is stored together with all the inferences about it. For the purposes of categorisation task in the current study DORA's retrieval, mapping, and schematisation routines were employed. The retrieval routine retrieves objects and relational structures from LTM based on the similarity of those representations to representations in AM. The mapping routine aligns objects and relational structures based on their similarities and roles within the relational structure supporting analogical inference. The schematisation routine refines the learned relational schemas, identifying common roles and relations they share and dismisses unshared features. Please see Doumas et al. (2008) for the more detailed description as well as for the model parameter values.

Following the interpretation of Goldwater and colleagues, one of the assumptions in the current study was that human participants' learning strategies differed in terms of using features versus relations and therefore could be captured by the difference in encodings. To simulate human participants employing featural and relational strategies on a category learning task, DORA was run with two different types of encodings. DORA with the featural type of encoding was limited to featural representation of objects and was prevented from creating a relational structure. Each stimulus was encoded as three lines with colour distributions serving as features. For instance, an example of a snarg in Figure 1a would be defined through featural encoding as follows:

line1: *b1, y1, g1, r1, r2, g2, g3, g4, r3, len9*
line2: *g5, g6, r4, y2, r5, b2, r6, r7, y3, g7, g8, r8, len12*
line3: *y4, r9, g9, g10, g11, g12, r10, r11, r12, b3, g13, r13, g14, r14, r15, g15, len16,*

where *r* stands for red square, *b* for blue, *g* for green, *y* for yellow, *len* for length.

Categorisation was performed with an exemplar approach. While various approaches to categorisation exist (see Pothos & Wills, 2011) the exemplar approach was chosen for two reason. First, it is in line with Goldwater et al.'s strategy in Experiment 2 of using the word-learning task to differentiate rule-learners from learners attempting to memorize exemplars. Second, the exemplar approach allowed us to employ DORA as a naïve category learner for the sake of simplicity—to investigate simplest possible categorisation mechanisms that do not involve learning.

The model was presented with an exemplar which was first placed in AM. DORA then attempted to retrieve one of the exemplars already stored in LTM (initially there were no exemplars in LTM) based on the featural similarity. Next, DORA attempted to map the current exemplar to any retrieved representations. If a mapping was discovered, then DORA categorised the current exemplar as the same category as the mapped exemplar. DORA then received feedback. The labeled exemplar entered LTM and could be retrieved in future categorisation trials.

With a relational encoding a stimulus was represented with a relational structure encoded as a ternary proposition (see Figure 2).

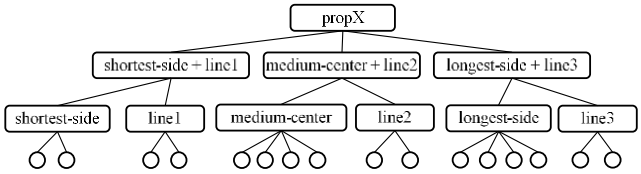

Figure 2: A ternary proposition which encodes the relational structure of a snarg shown in Figure 1a. Features are represented schematically by the lower-level circles.

A ternary proposition was utilised instead of four binary propositions since one proposition constitutes smaller cognitive load than four propositions. An example of four propositions in this case could be *shorter(line1, line2), shorter(line2, line3), left-of(line1, line2), and left-of(line2, line3)*. The use of ternary propositions is supported by empirical evidence demonstrating human capability to parse ternary relations as early as 5 years of age (Halford, 2014; Morrison et al., 2011).

A relational representation included three objects—three lines of one exemplar—and their roles in the relational structure which were assigned based on the relative line length and the position of the line in a stimulus (see Figure 2).

Importantly, the goal of the current study was to explore category learning rather than concept learning (e.g., Markman & Ross, 2003). A relational category is defined by a relational concept (e.g., the relational category of items of monotonically increasing size is defined by the concept of increasing size). As such, learning a relational category requires having the necessary relational concept. We assume that college students (human participants) are already familiar with concepts like increasing size and come into the study with these representations. During categorisation task what these participants are likely learning is to apply a known concept in a novel context. As such, we started DORA with representations of relations such as *larger* and *next*-to. For the current simulations the relations were hand-coded as learning these relations from examples has been previously simulated (e.g., Doumas et al., 2008).

Simulation 1 In the first simulation we simulated Experiment 1 from Goldwater et al. (2018). During categorisation task, DORA was run with either featural or relational encoding described above. The motivation for the original empirical experiment was to establish a proof of concept of relational versus featural strategy in category learning. The aim of Simulation 1 was to compare DORA's accuracy to human performance. If DORA's performance is comparable to that of human participants, this would offer a support to the notion that featural learners and rule-learners differ on the type of encoding engaged during category learning. We ran 50 feature-based simulations and 50 relation-based simulations.

Simulation 2 The second simulation simulated Experiment 2 of Goldwater et al. (2018). In Experiment 2 human participants were trained in either a blocking or an interleaving condition. These conditions focus on the sequence of stimuli presentation—how often stimuli from the same category are presented on the consequent trials. The chance of a stimulus from the same category being presented on the consequent trial is 75% in the blocking condition and 25% in the interleaving condition. We simulated these training regimes in DORA.

Simulation 3 In Experiment 2 accuracy of human participants on relation test trial was higher after blocked training compared to performance on relation test trial after interleaved training. Since DORA did not replicate this difference during Simulation 2, the third simulation introduced a recency bias in which DORA favoured the most recent exemplar for comparison with the current exemplar. This bias was instantiated in DORA as the ability to compare the current with the most recent exemplar—a direct comparison tactic previously shown to enhance category learning (Sandhofer & Doumas, 2008). The recency bias allowed us a preliminary investigation into whether the direct comparison is behind the relational advantage in humans after blocked training. While being trained in blocked and interleaved regimes with the recency bias, DORA compared the current exemplar with the most recent exemplar on the subject of category membership via mapping routine. Every time the direct comparison between two consecutive exemplars was relational and the classification was performed correctly, the model constructed a relational schema (see Doumas et al. 2008 for details of mapping and relational schematisation routines). The probability of the exemplar encoding to be relational on every trial was set at .5. This represented the idea that while humans do have relational representations they might not always use them during category learning.

Results

Simulation 1 Results and Discussion

During the learning and testing phases of Simulation 1 DORA demonstrated trends similar to those of human participants in Experiment 1 of Goldwater et al. (2018). During learning phase, as well as on baseline test trials, accuracy of both feature-based and relation-based strategies in human participants and of DORA running with feature-based and relation-based encodings were above chance. The relational strategy in humans and relational encoding in DORA produced slightly slower learning (see Figure 3).

Please note, in all figures, *"feature-based strategy"* and *"feature-based encoding"* are referred to as *"exemplar",* and "*relation-based strategy"* and *"relation-based encoding"* as *"rule"*).

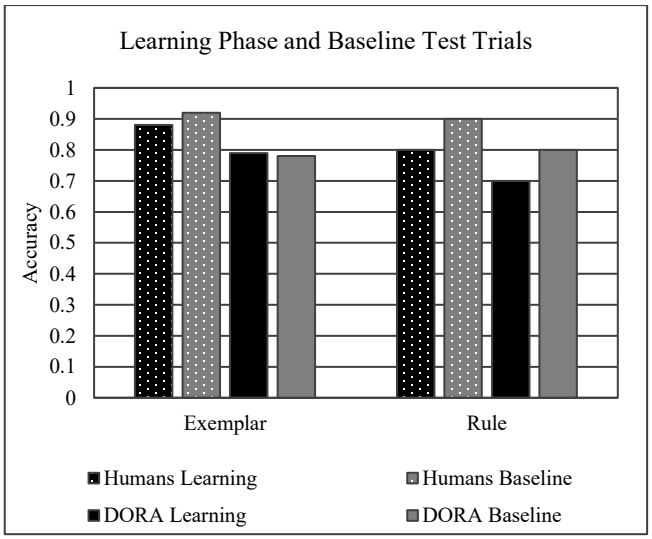

Figure 3: Learning phase and baseline test trials. Accuracy of exemplar versus rule approach in human participants in Experiment 1 and DORA's featural versus relational type of encoding in Simulation 1.

As noted above, on the feature test trials, stimuli were stripped of relational information by making all three lines in a stimulus of equal length. As expected, the accuracy of relation-based human participants was at chance, while those employing a feature-based strategy demonstrated high accuracy. Similarly, DORA run with relation-based

encodings on feature test trial performed close to chance, while feature-based encodings on the same kind of trial yielded high accuracy (Figure 4). This trend was reversed for the relation test trials in which distributions of four colours were equal. Feature-based human participants were near chance; by contrast, rule-learners performed with high accuracy. DORA demonstrated similar results, with featural encodings yielding near chance accuracy and relational encodings producing high accuracy (see Figure 4).

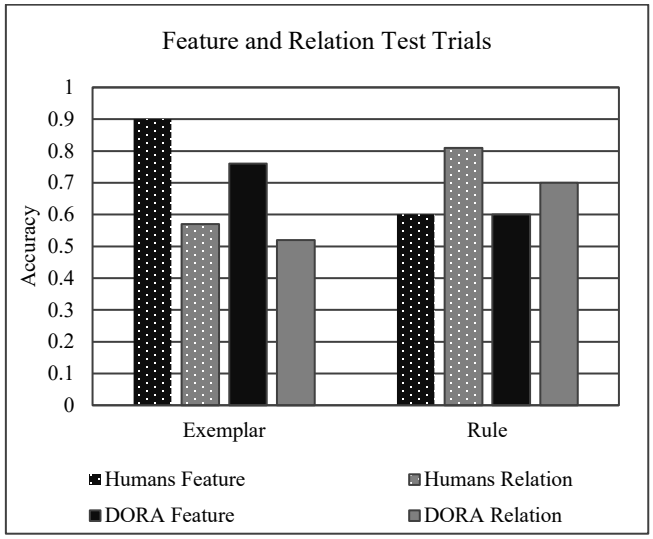

Figure 4: Feature and relation test trials. Accuracy of exemplar versus rule approach in human participants in Experiment 1 and DORA featural versus relational type of encoding in Simulation 1.

During cross-mapped test trials, the features and relations pointed at complementary categories. That is, on cross-mapped trials snargs had the prevalence of blue and yellow squares and blickets had the prevalence of reds and greens, while the opposite was true during training. The relations between lengths of the lines were preserved, with monotonically increasing or decreasing lines still defining snargs and non-monotonic relation defining blickets.

As expected, feature-based learners categorised using features, and thus performed systematically below chance when accuracy was based on the relational rule. Interestingly, the rule-based learners were not as accurate on cross-mapping trial as on baseline or relation trials, even though the rule did not change (Figure 5). DORA showed the same result with relational encoding. In DORA, the reduced accuracy was a product of the fact that relational representations included featural information, and categorisation was the result of retrieval from memory based on featural as well as relational similarity. Thus, featural information biased retrieval of category exemplars based on features rather than relational information. This result suggests that rule-learners might not be completely biased towards relational information, as it is often conceptualized in the literature, and might be lulled by featural similarity.

In short, Simulation 1 supported the notion that the differences between the strategies that human participants employ on the category learning task might be due to the differences in the types of stimulus encoding engaged during the categorisation task.

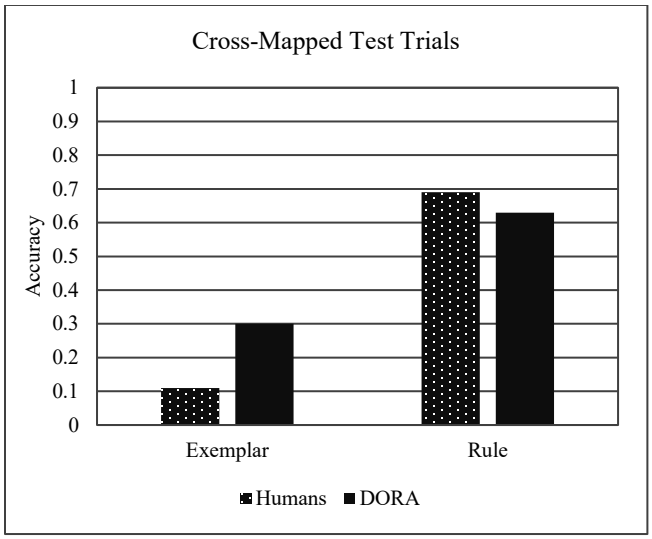

Figure 5: Cross-mapped test trials. Accuracy of exemplar versus rule approach in human participants in Experiment 1 and DORA's featural versus relational type of encoding in Simulation 1.

Simulation 2 Results and Discussion

In Goldwater et al.'s (2018) Experiment 2, rule learners trained with interleaving stimuli performed just as well as feature learners on feature test trials. That is, participants who reported using relational rules did not have trouble correctly categorising exemplars wherein the relational information was removed (Figure 6). Rule-learners trained with blocked trials, by contrast, did show the expected decrease in accuracy when relational information was removed (Figure 6). DORA showed the expected decrease in accuracy on feature trials when using relational encodings for both interleaved and blocked training. This result might suggest that differentiating between learning strategies with a self-report measure may not prove accurate.

On relation test trial DORA performed at chance when using featural encodings, and much better when using relational encodings (Figure 7). This result held after both types of training—with interleaved and blocked stimuli. The human participants showed this trend only with blocked training (Figure 7). That is, while feature learners performed close to chance on relation test trials with both interleaved and blocked training, rule learners performed well on relation test trials only with blocked training.

The same discrepancy between DORA and human participants was observed on the cross-mapped test trials (Figure 8). Both feature learners and DORA using featural encodings showed accuracy lower than chance—meaning that the opposite category was identified in most cases—

given both interleaved and blocked training. By contrast, humans who reported using a relation-based strategy performed below chance (i.e., used a the featural rule) when trained with interleaved stimuli, and at chance when trained with blocked stimuli. DORA did not show this trend, and rather showed feature-based reasoning (below chance accuracy) when using featural encodings, and relation-based reasoning (above chance accuracy) when using relational encodings.

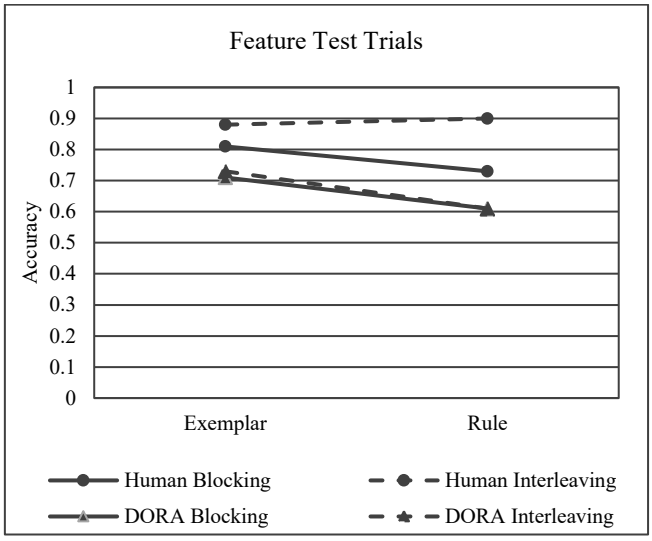

Figure 6: Feature test trials. Accuracy of exemplar versus rule approach in human participants in Experiment 2 and DORA's featural versus relational type of encoding in Simulation 2.

Goldwater et al. (2018) argued that the blocked training should benefit performance of the rule-based learners, while the interleaved training should enhance accuracy of feature-based learners. This prediction follows current thinking in analogy literature, i.e., relational category learning is promoted by analogical comparison (Christie & Gentner, 2010; Doumas et al., 2008), and analogy-making is easier when the compared exemplars are indeed from the same category as there are more similarities (relational or featural) to align.

Blocking versus interleaving affected neither accuracy during the learning phase, nor the baseline test trial performance for both, human participants and DORA. However, compared to the interleaving condition, human participants in blocking condition demonstrated higher accuracy on relation test trial. Since featural information was inaccessible, rule-learners performed better than exemplar-learners. It is interesting, however, that on the relation test trial in the interleaving condition human participants who reported rule-based learning strategy performed at chance. It was shown in literature that alternating between categories interferes with learning while presenting categories consequently enhances it (e.g., Sandhofer & Doumas, 2008). This pattern was not replicated in DORA and the model did not exhibit different behaviour on relation test trial after

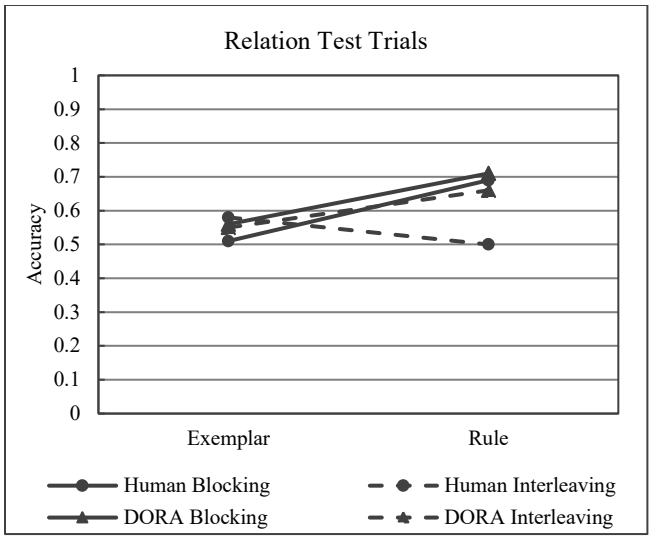

Figure 7: Relation test trials in blocking and interleaving conditions. Accuracy of exemplar versus rule approach in human participants in Experiment 2 and DORA's featural versus relational type of encoding in Simulation 2.

blocked versus interleaved training. One possible explanation is that the model in Simulation 2 was insensitive to the ordering of the exemplars. The goal of Simulation 3 was to address this limitation and to provide guidelines for future research.

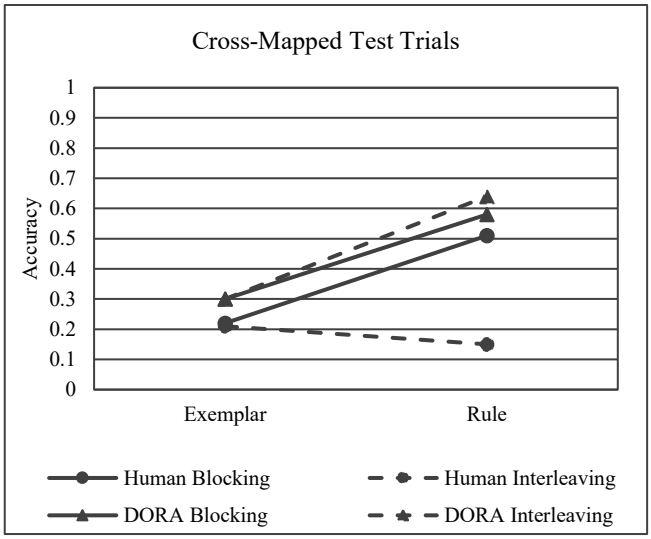

Figure 8: Cross-mapped test trials in blocking and interleaving conditions. Accuracy of exemplar- versus rule approach in human participants in Experiment 2 and DORA's featural versus relational type of encoding in Simulation 2.

Simulation 3 Results and Discussion

During Simulation 3 we implemented a recency bias in the model. DORA performs schematisation whenever it maps relational propositions. We instantiated the feedback in the

study by allowing the model to store schemas only after the classification was performed correctly. This resulted in different number of relational schemas after different training regimes. The blocked learning regime produced 32 relational schemas on average, while interleaved regime produced an average of 10 relational schemas after 20 simulations. These are reasonable results given that the probability of two consecutive exemplars being from the same category during blocked training is .75, and only .25 during interleaved training. Thus, after blocked training, the model has three times more schemas. These results suggest that indeed, as speculated by Goldwater et al. (2018), the advantage of blocked training was the invitation to compare more easily alignable items (i.e., items from the same category). By comparing items from the same category, assuming the model had noticed the correct relations, it could learn a schema defining a particular category. The resulting schemas were essentially variabalised representations of the key relational concept, and so made excellent representations of the relational category. Our simulations support the idea that a possible explanation for human participants' accuracy on relation test trial in blocking versus interleaving conditions is the generation of relational schemas defining the key relational category.

General Discussion

One goal of the computational simulations comprising this study was to investigate whether differences between feature-based and relation-based strategies employed by human participants in category learning task might be explained by the differences in the stimulus representational encoding mechanisms. Another goal was to test whether stimulus presentation sequence (blocking versus interleaving) affects performance differently when these types of encoding are engaged. We used DORA, a symbolic-connectionist model of relational learning and reasoning, to simulate performance on Experiments 1 and 2 of Goldwater et al. (2018) in which featural versus relational encodings were used to represent category exemplars.

Solely adjusting the representational encodings used by the model produced some striking similarities to human participants in Experiments 1 and 2 of Goldwater et al. (2018) in both accuracy and general trends. The results suggest that the differences between the exemplar and rule-based strategies employed by human participants during the learning stage might be due to the differences in the type of encoding engaged in the category learning task.

The results of human participants and DORA diverged in interesting ways in simulating Experiment 2. DORA, a model whose retrieval and mapping routines depend on analogical alignment, did not suffer any accuracy impairment given interleaved training compared to blocked training in Simulation 2. This outcome was unsurprising given that the difference between blocked and interleaved training is hypothesised to arise because blocked (as opposed to interleaved) training is said to promote comparison (e.g., Doumas & Hummel, 2013; Kurtz, Boukrina, & Gentner, 2013; Sandhofer & Doumas, 2008), which in turn promotes relation learning. The results of Simulation 3 suggest that direct comparison might influence the frequency of relational schemas construction which in turn could be the mechanism behind the relational category learning. It has been demonstrated in analogy research that comparison facilitates learning relational schemas (e.g., Doumas & Hummel, 2013; Gick & Holyoak, 1983; Hummel & Holyoak, 1997; Jamrozik & Gentner, 2013; Kurtz et al., 2013). Future simulations should explore whether the representational structure of relational category corresponds to the relational schema in analogy research or whether relational representations of relational categories are governed by the same mechanisms as featural representations.

Finally, consider the drop in accuracy for the rule-based learners and relation-based DORA on cross-mapped trials. This trend occurred despite the fact that while feature-based categories were swapped on these trials, the relational rule did not change. Rule-learners performance at chance on cross-mapped test trial suggests that they were affected by the featural information and did not necessarily focus purely on relational information. DORA's performance also suffered. DORA was affected by the feature swap (Figure 8) as object features are present even in the relational encoding. This is an interesting result which suggests the reason why human relational categorisation might remain affected by featural similarity.

The current computational study has several limitations. One of them being the fact that model's LTM was empty prior to the simulations. An empty LTM means that DORA had no prior knowledge whatsoever before the task. Future simulations need to address this issue by pre-training DORA on other data containing spatial information to learn relations such as *longer (a, b), shorter(c, d), to-the-side(e, f), and in-the-middle(g, h)*.

One of the limitations of LISA/DORA models is the lack of feature salience. Since feature salience affects performance as demonstrated in the Experiment 3 in Goldwater et al. (2018), this aspect needs to be added to the model in the future.

In summary, the results of the simulations suggest that feature- and relation-based strategies that human participants employ on categorisation tasks might be the result of engaging different representational encodings—by encoding each stimulus purely with its features or with a relational structure. In addition, our findings suggest that within-category comparisons during training enhance relational categorisation performance. This enhancement might be due to the fact that direct comparison of relationally similar consequent elements in blocked training facilitates generation of relational schemas which define a relational category.

Acknowledgments

The authors would like to thank three anonymous reviewers for helpful comments and suggestions and Dr. Micah Goldwater for providing experimental results.

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87-113.

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, *11*(3), 356-373.

Corral, D., Kurtz, K. J., & Jones, M. (2018). Learning relational concepts from within-versus between-category comparisons. *Journal of Experimental Psychology: General*, *147*(11), 1571-1596.

Davis, T., Goldwater, M., & Giron, J. (2017). From concrete examples to abstract relations: The rostrolateral prefrontal cortex integrates novel examples into relational categories. *Cerebral Cortex*, *27*(4), 2652-2670.

Doumas, L. A., & Hummel, J. E. (2005). A symbolic-connectionist model of relation discovery. In *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 27, No. 27).

Doumas, L. A., & Hummel, J. E. (2010). A computational account of the development of the generalization of shape information. *Cognitive Science*, *34*(4), 698-712.

Doumas, L. A., & Hummel, J. E. (2013). Comparison and mapping facilitate relation discovery and predication. *PloS one*, *8*(6).

Doumas, L. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, *115*(1), 1-43.

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

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

Goldwater, M. B., Don, H. J., Krusche, M. J., & Livesey, E. J. (2018). Relational discovery in category learning. *Journal of Experimental Psychology: General*, *147*(1), 1-35.

Halford, G. S. (2014). *Children's understanding: The development of mental models*. Psychology Press.

Higgins, E., & Ross, B. (2011). Comparisons in category learning: How best to compare for what. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 33, No. 33).

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*(3), 427-466.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*(2), 220-264.

Jamrozik, A., & Gentner, D. (2013). Relational labels can improve relational retrieval. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 651-656).

Jung, W., & Hummel, J. E. (2015). Making probabilistic relational categories learnable. *Cognitive Science*, *39*(6), 1259-1291.

Jung, W., & Hummel, J. (2013). The effects of dual verbal and visual tasks on featural vs. relational category learning. *In Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, No. 35).

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*(4), 1303-1310.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological bulletin*, *129*(4), 592-613.

Martin, A. E., & Doumas, L. A. (2019). Predicate learning in neural systems: Using oscillations to discover latent structure. *Current Opinion in Behavioral Sciences*, *29*, 77-83.

Morrison, R. G., Doumas, L. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: Balancing inhibitory control in working memory and relational representation. *Developmental Science*, *14*(3), 516-529.

Pothos, E. M., & Wills, A. J. (Eds.). (2011). *Formal approaches in categorization*. Cambridge University Press.

Sandhofer, C. M., & Doumas, L. A. (2008). Order of presentation effects in learning colour categories. *Journal of Cognition and Development*, *9*(2), 194-221.