Title

XOR in Order: Category Learning of Exclusive-Or in a Temporal Sequence

Permalink

https://escholarship.org/uc/item/14r1p3hh

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Shah, Vedant Biren
Schlegelmilch, René
von Helversen, Bettina

Publication Date

2024

Peer reviewed

XOR in Order: Category Learning of Exclusive-Or in a Temporal Sequence

Vedant Shah (vshah@uni-bremen.de)
Department of Psychology, University of Bremen, GER

René Schlegelmilch
Department of Psychology, University of Bremen, GER

Bettina von Helversen
Department of Psychology, University of Bremen, GER

Abstract

When people make decisions, these often do not stand alone but are made in a sequence of decisions. For instance, a doctor will first decide on a patient's treatment and then about the duration of the treatment. In such decision sequences, later decisions frequently depend on the outcome of the first decision (e.g., if the treatment results in an adverse reaction, this predicts the next decision). While there is research on how humans discover inter-relations between sequentially presented information, for example, regarding grammar, little is known about whether and how humans can also learn complex inter-category relations in decision sequences. To provide a step towards closing this gap, we present an experiment in which we embedded a problem structure, known in category learning as Type II or Exclusive-Or, in a sequence of three decisions. In each trial, participants saw one of eight unique stimuli, each followed by three categorization tasks for this stimulus. In a Type II condition, the outcomes of tasks 1 and 2 predicted the outcome of task 3, which we compared to a control condition without a regularity. We hypothesized that the sequential Type II regularity, as in visual category learning, would facilitate learning and subsequent generalization compared to the control condition. Instead, the evidence favored the Null hypothesis in both cases. This is in contrast to findings from the visual categorization domain in which this benefit is reliably observed. These findings highlight the boundary constraints on the human ability to discover rule-based category structures in sequential sequences.

Keywords: Category Learning; Sequential Learning

Introduction

Human learning often relies on discerning patterns and regularities in the surrounding environment, as those usually follow specific rules, as studied in language, music, motor sequencing, visuospatial perception (Gomez, 2002; Newport et al., 2004b,a; Romberg & Saffran, 2013; Vuong et al., 2016; Lu & Mintz, 2021; Iao et al., 2021), and category learning (e.g., Shepard et al., 1961; Nosofsky, Gluck, et al., 1994). In category learning, it is typical to study the corresponding learning processes using visual stimuli that differ on multiple visual dimensions (e.g., color, size). Participants have to predict which category a stimulus belongs to based on the visual dimensions (e.g., green things are bananas, blue things are berries). However, it is plausible that humans can also learn inter-category relations, which are surprisingly rarely studied. Consequently, we extend the classic approach to category learning by investigating whether learning phenomena observed in visual scenarios generalize to situations where a predictable structure arises from correlated categories in a sequence of stimulus-category decisions.

For example, imagine you are inspecting some rocks based on their visual properties. First, you categorize the rock as containing metal or not. In the next step, you categorize it as heavy or light. While both could be done based on the rock's appearance, you could notice that metal rocks are also heavy after examining many rocks. Consequently, you can predict its weight (heavy vs. light) once you know the first outcome (metal or not) without requiring visual information.

The current work focuses on better understanding under which conditions people can learn to use such sequential inter-category associations, as illustrated in the rock example above, and how this compares to how people use visual stimulus dimensions in category learning. Visual category learning has shown that the ease with which people master a visual categorization task depends on the category structure, as illustrated in the seminal work by Shepard et al. (1961)(see also Nosofsky, Gluck, et al., 1994; Lewandowsky, 2011). In this classic work, Shepard et al. (1961) introduced six problem types in which participants categorize stimuli based on three binary dimensions (i.e., color, size, and shape) into two categories. Figure 1 demonstrates three of these problems. In Type I problems, a single binary feature is sufficient to predict category membership (i.e., large objects belong to Category A and small objects to Category B). Type II problems, a.k.a. Exclusive-Or (XOR), can be solved by a combination of two binary features (e.g., size [small vs. large], shape [triangle vs. square]), whereas the Type VI problem is unstructured, and predicting the categories requires memorizing every stimulus and its category.

The above rock example would correspond to the Type I problem, which seems straight forward. However, sequential-intercategory relations can also be conceived as Type II problem when they include conditional regularities. As an everyday example, consider a security mechanism of a door, with sensors recognizing whether the door is open or not, and whether a correct security code was entered or not. The third decision is regarding whether the security system works correctly. The decision that the system works correctly follows from two cases, if the door is open while somebody entered a correct security code, or when the door is closed while no code was entered. Vice versa, a decision that the system does not work follows from the opposite cases, namely, if the door is open while no code was entered, or if the door is closed despite the correct code being entered. Returning to our rock

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

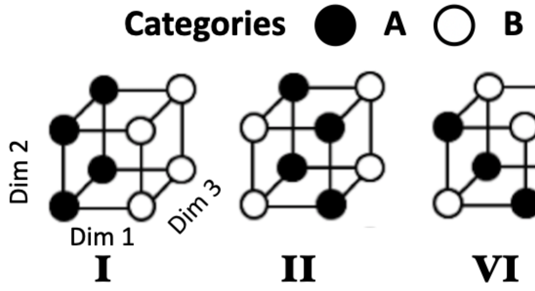

Figure 1: Category Structures. Type I, II and VI from Shepard et al. (1961). Coordinates reflect three binary dimensions (e.g., Dim 1 = size: large vs. small). Shading reflects assigned category (A vs. B). See further text.

example, a decision might concern, whether a rock is to-be further 'processed' or 'not'. In a Type II combination this decision could depend on the rock being metal or not, and heavy or not. For instance, if 'metal' and 'heavy' or 'non-metal' and 'light' further categorize the rock as to-be-processed, but if 'metal' and 'light' or 'non-metal' and 'heavy' further categorize the rock as not-to-be-processed.

In visual categorization, research suggests that people learn Type I problems much more quickly than Type VI problems. Similarly, the XOR (Type II) structure is often found to be easier than problems requiring to memorize the stimulus-category association as in Type VI (e.g., Shepard et al., 1961; Nosofsky, Gluck, et al., 1994; Lewandowsky, 2011) but more difficult compared to Type I. These differences in learning difficulty are often attributed to using different cognitive processes. In Type I and II problems, people likely realize that they can abstract a perfectly predictive categorization rule using just one or two features. In contrast, in Type VI problems, people have to memorize each exemplar and its associated category, a slower and more effortful process compared to Type I and Type II. This provides support concerning different learning strategies, such as rule vs. memory-based processing (Anderson, 1991; Nosofsky, 1991a,b; Nosofsky, Palmeri, & McKinley, 1994; Schlegelmilch et al., 2021; Bruner et al., 1956; Martin & Caramazza, 1980).

However, it is an open question whether the same phenomenon will occur when these regularities are observed in a temporal sequence of categorical outcomes (inter-category instead of feature-category), like the rock example. The focus here is on the categorical outcome of each stimulus and not on the visual features of the stimuli presented. In our previous research (https://osf.io/gd9v4/), we therefore first investigated whether people can learn a temporal version of the Type I category structure. For this, we presented one of multiple idiosyncratic stimuli (e.g., an icon representing a mountain or a knot), for which participants made two subsequent categorical outcome predictions (T1: A vs. B, then T2: C vs. D). While there was no visual regularity to predict these categories, we embedded a Type I rule in the two-category sequence (i.e., T1: A always predicted T2: C, and T1: B always predicted T2: D). In other words, participants could either use the visual stimulus to memorize its category, or they could pick up on the Type I between-category relation. We compared this task with a Type VI structure (control condition), which participants could only master by memorizing all stimulus-category associations. Crucially, after several learning trials, we also introduced novel stimuli but withholding feedback to test whether participants could generalize the Type I regularity. The left panel of Figure 2 shows the learning curves for the critical outcome T2 in both conditions, averaged within the ten repeated training blocks. As can be seen, T2 accuracy was higher in the Type I condition compared to the control Type VI condition. Furthermore, the right panel of Figure 2, shows the probability of regularity-consistent responding for novel items in the generalization test phase. As can be seen, we found conclusive evidence ($BF_{10} > 100$; M = 75.33, $CI95$= [65.98, 84.69]), indicating that participants successfully generalized the Type I regularity to novel objects.

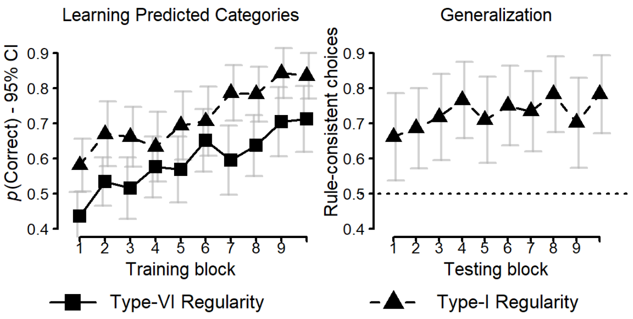

Figure 2: Experiment Type I vs Type VI. The left panel shows accurancy (y-axis) over training blocks (x-axis; 8 trials per block) for T2 categorizations. The right panel shows the proportion of regularity-consistent choices (y-axis) over test blocks (x-axis; 4 trials per block) for T2. Error bars indicate 95% CIs of individual means.

From this perspective, people might be able not only to learn simple associations (e.g., Type I) but also more complex ones (XOR), when embedded into sequential decisions compared to a Type VI problem, which we test in the current study. Indeed, similar questions on the influence of structural complexity in sequential learning have been investigated in the domain of grammar learning (see; Gomez, 2002; Romberg & Saffran, 2013; Deocampo et al., 2019; Newport et al., 2004b), in which sequentially observed grammatical elements predicted further elements, for example, syllables or sounds in a tri-element sequence (e.g., ba→da and pa→do in a stream of ba-da-ku, pa-do-ti, ba-da-ti, etc.). First, these studies suggest that people could learn such tri-element syllables with inter-element regularity better than tri-elements without regularity, which also leads to generalization for novel tri-element sequences. Second, however, the ease of acquisition, in terms of complexity, seems to co-depend on whether these regularities appear in adjacent vs. non-adjacent (element regularities with an intervening element) positions (e.g., Newport et al., 2004b; Conway & Christiansen, 2005; Wilson et al., 2020), and/or where the adjacent and non-adjacent associations occur concurrently (e.g., Deocampo et al., 2019; Conway et al., 2020, e.g., when ba→da and pa→do, the presented sequence is either ba-pa-da-do (crossed regularity) or ba-pa-do-da (nested regularity)).

These insights from grammar learning seem to suggest that it might be possible to learn complex sequential regularities like XOR in decision sequences as well. Formally, consider a sequence of three categorical decisions for the same stimulus (T1: A vs. B, T2: C vs. D, and T3: E vs. F). A temporal XOR structure here means that the first *two* categories (T1 & T2) disjunctively predict the third one (T3), such that AC→E and BD→E but AD→F and BC→F. However, there might be reasons to expect no learning of XOR regularities. For one, in the visual domain, Love & Markman (2003) suggests that changing the stimulus material might hinder discovering complex regularities if the stimulus features were perceived as less separable, or as highlighted by Kurtz et al. (2012), if participants are not directly prompted to search for categorization rules. From another perspective, when compared to complex grammar learning, another reason might be that discovering XOR structures in decision sequences is more difficult than rule discovery in rapidly presented phonological or syntactical elements. That is, an XOR solution in a sequential categorization task requires keeping category information in mind over successive actions, which seems more demanding than combining concurrently presented visual features. Thus, answering the question whether or not XOR can be learned in a sequence of decisions could further shed light on the boundary constraints of rule discovery in category and sequential learning in general.

Experiment

The experiment was designed to find out whether participants can learn and generalize disjunctive rules known as Type II or XOR when embedded in a sequence of decisions. Our task design reflects a trial-and-feedback learning procedure. In each trial, participants see an idiosyncratic stimulus followed by three sequential binary categorization tasks. For each feedback is provided after the corresponding decision. In the following we therefore refer to the inter-category dependencies in terms of task 1, 2 and 3 (for T1, T2, and T3, respectively). In the Type VI condition, the outcomes of each task were independent of each other. In the Type II condition, the outcomes of task 1 and 2 predicted the outcome of task 3 in an XOR structure. The study was pre-registered on OSF (https://osf.io/jgwhr).

Method

Participants As pre-registered, we implemented the Sequential Bayes factor method with maximal N design (Schönbrodt & Wagenmakers, 2018), taking the hypothesis test on the assumed interaction in the test phase for novel stimuli (generalization, see Results section). That is, we determined a minimum (N_{min}) = 60 participants (i.e. 30 per condition) and continued the sampling until a BF>10 or BF<0.1 was reached. For determining N_{max}, we simulated the category output in the Type II condition according to a small effect p = .6 rule-consistent responding in the Type II condition (i.e., the mean probability of the consistent response based on a Binomial distribution for each individual in 16 trials [4 novel stimuli * 4 blocks]), then conducted a one-sample Bayesian t-test on the simulated data against mean = .5. We repeated the procedure 500 times, with different sample sizes, and obtained a BF>10 in 90.4%, with 60 participants in each condition, which we set as N_{max} (https://osf.io/mzkyt).

We recruited 70 English-speaking participants on Prolific Academic (Male = 42, Female = 28; *M*(age) = 32, *SD* = 7). They received 4.00€ as basic compensation and an additional bonus, which increased with their learning accuracy (2€ max). The participants were randomly assigned to the two conditions. The study duration was *M* =32.8 minutes, *SD* = 16.5 minutes. To ensure equal levels of memory performance, we excluded 6 participants based on the method explained below (N(Type VI)=5, N(Type II)=1) for not learning task 1. One participant indicated that their data should not be used for the analysis, which we also excluded. Overall, 30 and 33 participants remained in the Type VI and Type II conditions, respectively.

Design The learning task consisted of a sequence of three sub-tasks on fictitious objects (see Table 1), similar to our initial rock-analysis example. Task 1, 2 , and 3 comprised categorization tasks, which had to be performed for each object. In each task, participants had to classify the object into one of two categories (T1: Herf vs. Jonth; T2: Krill vs. Wask; T3: Thesh vs. Aurek). There were two experimental conditions: In the experimental condition, also referred to as the Type II condition, the outcomes of task 1 and 2 predicted that of task 3, regardless of the assigned visual object, as in the example in Table 1. In the control condition, also referred to as Type VI condition, there was no dependency between the three tasks. Thus, the study implements a 2 (Type II regularity vs. control; between) x 3 (Task 1 vs. 2 vs. 3; within) mixed factorial design.

Material and Procedure The experiment was created using jsPsych De Leeuw (2015) and conducted online on Prolific. The experiment comprised two phases: training and test. During the training phase, participants engaged in three decision tasks using a trial-and-feedback approach, repeatedly encountering eight objects in a block-wise manner twelve times. The object presentation order within blocks was randomized, resulting in 96 trials. The eight objects were chosen randomly for each participant from a set of 12 objects (created by Freepik) and assigned to the category structure. One random

Table 1: **Task Design - Type II Regularity**

Object	Task 1	Task 2	Task 3
	Herf	Krill	Aurek
	Herf	Krill	Aurek
	Herf	Wask	Thesh
	Herf	Wask	Thesh
	Jonth	Wask	Aurek
	Jonth	Wask	Aurek
	Jonth	Krill	Thesh
	Jonth	Krill	Thesh

Task 1-3 refer to subsequent decisions for the same object. 'Herf' and 'Jonth' refer to category outcomes in Task 1, 'Krill' and 'Wask' to outcomes in Task 2, and 'Thesh' and 'Aurek' to outcomes in Task 3, with Type II structure to predict Task 3.

allocation is illustrated in Table 1. The remaining four objects served as novel objects for the test phase. Object-task assignments remained constant within participants across all trials but varied randomly between participants. In each trial, participants saw an object and then could categorize it by clicking on one of two buttons with the category labels that were shown beneath the object. The three categorization tasks followed sequentially until the third task was performed. Afterwards the next trial started. The button-label assignments were randomly interchanged over trials to emphasize learning category labels rather than specific button-press associations such as learning a sequence of pressing right-left-right for an object.

Participants followed an identical procedure in the test phase except for two changes. First, we showed four randomly selected objects from the training phase (one from each unique combination of the T1 and T2 category combinations). In addition, participants saw four novel objects, which we used to test whether participants made decisions consistent with Type II structure in the experimental condition. Second, participants only received feedback for tasks 1 and 2, but not for task 3, to cleanly asses rule generalization without ongoing learning. The test phase consisted of 32 trials (eight objects randomized within four blocks).

Before participants started the categorization tasks, they provided informed consent and received instructions. Before the training phase, we asked three control questions to ensure an understanding of the procedure. After the experiment, we asked participants to rate the perceived difficulty of the study and their diligence while participating as well as indicate any strategies they used while doing the tasks (not reported). We also asked participants if there were reasons to exclude their data from the analysis (e.g., use of tools). The participants who withdrew their consent were excluded from the analysis. Lastly, they provided the prolific code as proof of completion.

Data Cleansing To control for sample biases in overall performance we excluded participants with chance or lower-than-chance performance in the last 24 trials of the training blocks in task 1. For this, we pre-registered and used a Bayesian latent class model to classify participants into guessing, medium, and high-accuracy groups in the final three blocks of the training phase (24 decisions), which we performed before applying any hypothesis testing. Priors were set for each group mean probability as, respectively, $\phi_{guessing} = .5$, and ϕ_{medium} and ϕ_{high}, with the latter two drawn from a uniform distribution between .5 and 1 (non-hierarchical). The models were assigned to the participants in a trans-dimensional MCMC method drawing the model likelihoods from a Dirichlet distribution (uniform), passed to the individual level via categorical samples (Schlegelmilch & von Helversen, 2020; Zeigenfuse & Lee, 2010, for similar applications, see). We ran 20000 iterations to calculate how often each participant was assigned to each group and excluded the participants assigned to the guessing group with more than 80% confidence.

Results

Training Accuracy Our hypothesis was that learning would be better in a classical Type II structure when embedded in a sequence compared to a decision sequence without any structure (control). We performed a mixed-effects logistic regression on decision accuracy (1=correct) with training blocks, condition, and task as fixed effects, including by-participant and by-stimulus as random intercepts (using the R package Afex Singmann et al., 2015) with Type III LRT tests (full vs. reduced model). As also suggested in the Figure 3, we found no significant main effects for task, χ^2(1, 62) = 1.38, p = .503, or condition, (X^2 (1, 62) = 0.47, p = .495, suggesting no difference in accuracy in the three tasks or between the conditions. There was a significant main effect of blocks (X^2 (1, 62) = 15.52, $p < .001$), showing that accuracy increased with learning. Importantly, there was no significant interaction between condition and task, X^2(1, 62) = 2.03, p = .362, indicating that T3 accuracy, indeed, did not differ between the Type II and the control condition. There was also no interaction between task and blocks, X^2(1, 62) = 5.08, p = .079, condition and blocks, X^2(1, 62) = 1.59, p = .208, or a 3-way interaction, X^2(1, 62) = 2.33, p = .312.

We carried out planned post-hoc analyses on the mixed model estimates (log-scale; R package: Emmeans; Lenth et al., 2019), which further confirmed that participants in the Type II condition were not better in learning the outcome of task 3. Within participants, there was higher accuracy in task T1 compared to T3, $z = 3.22$, $p = .004$, *Mdiff* = 0.05, *CI95* = [0.048,0.3], however, no difference between T3 and T2, $z = 1.07$, $p = .53$, *Mdiff* = 0.05, *CI95* = [-0.07,0.19] in the Type II condition. In the Type VI condition, there was neither a difference in accuracy in task T3 compared to T1, $z = 1.17$, p = .47, *Mdiff* = 0.05, *CI95* = [-0.07,0.20], nor compared to

T2, $z = 0.6$, $p = .82$, *Mdiff* = 0.05, *CI95* = [-0.10,0.17]. Finally, there was no significant difference found in the learning accuracy in T3 between Type II and control (Type VI) conditions, z = -0.142, p = 0.8868, *Mdiff* = 0.2, *CI95* = [-0.42,0.36]. Thus, the training results suggest that learning Type II vs. Type VI (control) does not reproduce the ordinal difficulty trends observed in visual category learning. In other words, learning performance on the critical task 3 was equal between both conditions.

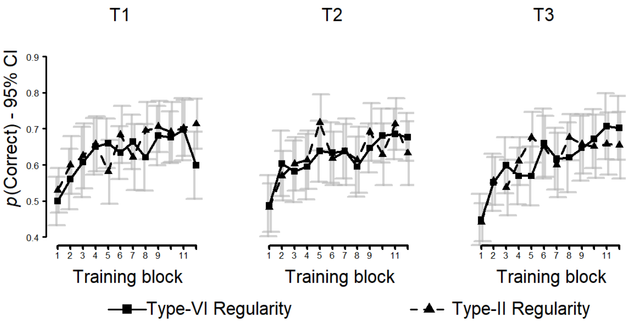

Figure 3: Training phase mean accuracy. Accuracy (y-axes) over training blocks (x-axes; 8 trials each block) for T1 categorizations (left), T2 (middle), and T3(left). Error bars indicate 95% CIs of individual means.

Testing Accuracy: Old Items We hypothesized that *previously trained* (old) stimuli were categorized equally well or better for task 3 in the Type II condition compared to the Type VI control condition. Figure 4 depicts the main results, again suggesting equal performance between sub-tasks and conditions. We tested this by performing the same mixed-effects logistic regression model as the one performed for the training phase but focusing on the four previously trained stimuli. We found no significant main effect of task, $X^2(1, 62) = 1.38$, p = .502), blocks,($X^2(1, 62)$ = 0.25, p = .62, and condition $X^2(1, 62)$ = 0.02, p = .85. There was no significant interaction between condition and blocks ($X^2(1, 62)$ = 0.01, p = .92, or between tasks and blocks ($X^2(1, 62)$ = 1.22, p = .55 or between task and condition ($X^2(1, 62)$ = 0.73, p = .69 or a three-way interaction, $X^2(1, 62)$ = 0.25, p = .88.

A planned post-hoc analysis on the interaction between outcome and condition (as pre-registered) further confirmed that there was no difference in the outcome accuracy of T3 in the Type II condition compared to the control condition, z = 0.21, p = .84, *Mdiff* = 0.35, *CI95* = [-0.61,0.76]. Thus, the test phase mainly replicates the previous results from the training phase.

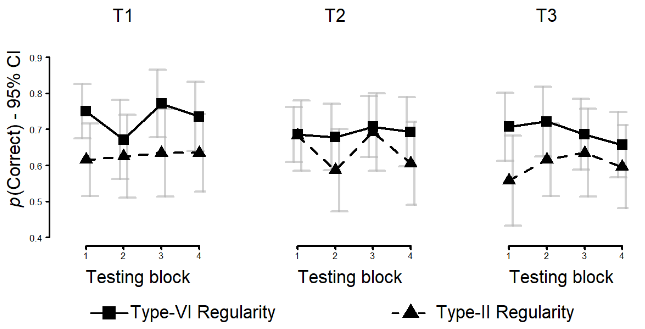

Figure 4: Test phase mean accuracy (old objects). Accuracy (y-axes) over testing blocks (x-axes; 4 trials each block) for T1 categorizations (left), T2 (middle) and T3 (left). Error bars indicate 95% CIs of individual means.

Testing Accuracy: Novel Items Finally, to see whether there nonetheless was some generalization of the Type II regularity on *novel* objects, we tested whether the participants' predictions were consistent with Type II. This would be indicated by higher-than-chance rule-consistent responding (above 50%) for these stimuli (recall, T1 and T2 were known via feedback, but feedback in task 3 was withheld). Consistent predictions are T3 choices that follow the same pattern as for old objects based on the observed outcomes of task 1 and 2. Figure 5 (right panel) depicts the average consistency score over the test blocks for the Type II condition, as done for our previous study on Type I (https://osf.io/gd9v4/). Since there is no regularity in the control condition, we tested the hypothesis focusing on Type II, using a one-sample Bayesian t-test against 50% (guessing). We found conclusive evidence against the hypothesis that participants generalized the Type II regularity, $BF_{10} = 0.21$. Thus, the generalization result is consistent with the results from the training phase. We therefore conclude, that embedding a Type II category structure in a sequence of decisions, unlike in visual category learning, does not facilitate learning compared to Type VI (control), further discussed below.

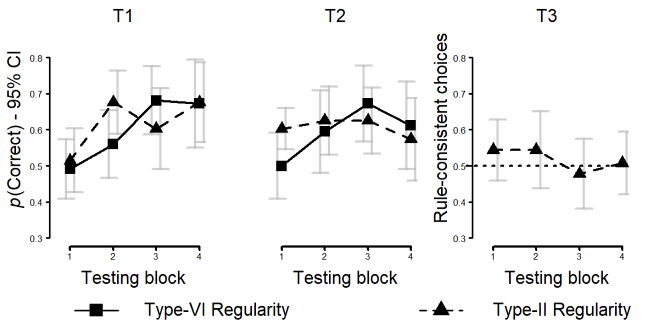

Figure 5: Testing phase (new targets). Accuracy (y-axes) over testing blocks (x-axes; 4 trials each block) for T1 categorizations (left-panel) and T2 (middle-panel). The right panel shows the propotion of rule-consistent categorizations (y-axes) over training blocks (x-axes; 4 trials each block) for T3 categorizations against chance level (50% accuracy). Error bars indicate 95% CIs of individual means.

Discussion

Our goal was to investigate whether humans can learn an XOR regularity, as traditionally observed in visual category learning, when embedded in three sequential categorization tasks in terms of an inter-category relation, and whether they generalize this regularity to novel stimuli. The results suggest that participants neither acquired nor generalized the sequential XOR structure. Indeed, the participants' performance was basically the same as in the control condition without inter-category regularity, in which each stimulus-category association had to be memorized. These results are inconsistent with visual category learning, where people perform better in tasks with an XOR structure than in Type IV tasks (e.g., Shepard et al., 1961; Nosofsky, Gluck, et al., 1994; Lewandowsky, 2011), and highlights a corresponding boundary condition. Below, we discuss some potential reasons and further implications.

First, our previous study showed that participants could acquire a Type I regularity in a sequence of decisions in a nearly identical procedure. Thus, the current lack of a benefit on learning and generalization is not solely due to the inability to pick up on temporal category regularities. The complete lack of a benefit seems also a little surprising in light of grammar learning studies showing that participants can learn more complex regularities at least to some degree (Gomez, 2002; Romberg & Saffran, 2013; Deocampo et al., 2019; Newport et al., 2004b). What differentiates the current Type II problem from such grammar learning studies is that those typically entail the mere observation of sequentially presented elements (e.g., sounds) in rapid succession (e.g., incidental learning), while our task included active decision-making with each decision taking some time. In other words, in each of the three tasks, participants had to integrate information to form a category prediction, which would recruit processes of category inference, while requiring to keep relevant information in short-term memory.

In a similar vein, previous studies on visual category learning suggest that not only the stimulus material can affect discovering XOR rules (Love & Markman, 2003), but also task instructions regarding whether or not to search for rules at the outset (Kurtz et al., 2012). Thus, a corresponding follow-up question could be under which circumstances participants also could learn a sequential XOR rule. It seems plausible, that instructions highlighting the existence of the inter-category regularities might boost learning and generalization performance. However, even if participants knew about the existence of a rule, noticing complex inter-category regularities like XOR necessitates good memory of both outcomes of tasks 1 and 2 while predicting the outcome of task 3. Thus, in contrast to grammar learning making decisions could interfere with maintaining this information nonetheless. It seems worthwhile investigating whether working memory capacity predicts rule discovery, while making the acquisition of the inter-category regularity intentional vs. incidental, highlighting the potential of extending such classic category learning designs to the temporal domain.

In particular, it seems fruitful to study problems such the sequential XOR tasks in relation to measures of complex working memory span (e.g., remembering stimuli, while performing other mental operations; see Redick & Lindsey, 2013). As other researchers highlighted, working memory is an inherent component of category learning (e.g., updating and maintaining of information), however, remains still an understudied topic (e.g., Lewandowsky, 2011). Thus, we believe that despite the Null result, our study provides a first step towards understanding the limitations of when and how humans acquire inter-category relations, and provides an outlook how these limitations could be overcome also in relation to individual differences in working memory. For instance, learning about compound predictors could be influenced by keeping more or less outcome information of previous tasks visually available (similar to set size in working memory studies), and varying the amount of information that needs to be maintained or integrated. Such studies could betray the interactions of multiple cognitive processes, which we seek to address in future studies.

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*(3), 409–429. doi:https://doi.org/10.1037/0033-295X.98.3.409

Bruner, J. S., Goodnow, J. J., & George, A. (1956). A study of thinking. *New York: John Wiley & Sons, Inc*, *14*, 330. doi:http://www.jstor.org/stable/27538750

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 24–39. doi:https://doi.org/10.1037/0278-7393.31.1.24

Conway, C. M., Eghbalzad, L., Deocampo, J. A., Smith, G. N., Na, S., & King, T. Z. (2020). Distinct neural networks for detecting violations of adjacent versus nonadjacent sequential dependencies: An fmri study. *Neurobiology of Learning and Memory*, *169*, 107175. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1074742720300198` doi:https://doi.org/10.1016/j.nlm.2020.107175

De Leeuw, J. R. (2015). jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*, 1–12. doi:https://doi.org/10.3758/s13428-014-0458-y

Deocampo, J. A., King, T. Z., & Conway, C. M. (2019). Concurrent learning of adjacent and nonadjacent dependencies in visuo-spatial and visuo-verbal sequences. *Frontiers in Psychology*, *10*, 1107. doi:https://doi.org/10.3389/fpsyg.2019.01107

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436. doi:https://doi.org/10.1111/1467-9280.00476

Iao, L.-S., Roeser, J., Justice, L., & Jones, G. (2021). Concurrent visual learning of adjacent and nonadjacent dependencies in adults and children. *Developmental Psychology*, *57*(5), 733–748. doi:https://doi.org/10.1037/dev0000998

Kurtz, K., Levering, K., Stanton, R., Romero, J., & Morris, S. (2012, 07). Human learning of elemental category structures: Revising the classic result of shepard, hovland, and jenkins (1961). *Journal of experimental psychology. Learning, memory, and cognition*, *39*. doi:10.1037/a0029178

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *Package 'emmeans'*.

Lewandowsky, S. (2011). Working memory capacity and categorization: Individual differences and modeling. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *37*(3), 720–738. doi:10.1037/a0022639

Love, B., & Markman, A. (2003, 05). The non-independence of stimulus properties in human.

Lu, H. S., & Mintz, T. H. (2021). Learning non-adjacent rules and non-adjacent dependencies from human actions in 9-month-old infants. *Plos One*, *16*(6), e0252959. doi:https://doi.org/10.1371/journal.pone.0252959

Martin, R. C., & Caramazza, A. (1980). Classification in well-defined and ill-defined categories: evidence for common processing strategies. *Journal of Experimental Psychology: General*, *109*(3), 320–353. doi:https://doi.org/10.1037/0096-3445.109.3.320

Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004a). Learning at a distance ii. statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, *49*(2), 85-117. doi:https://doi.org/10.1016/j.cogpsych.2003.12.002

Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004b). Learning at a distance i. statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. doi:https://doi.org/10.1016/S0010-0285(03)00128-2

Nosofsky, R. M. (1991a). Relation between the rational model and the context model of categorization. *Psychological Science*, *2*(6), 416–421. doi:https://doi.org/10.1111/j.1467-9280.1991.tb00176.x

Nosofsky, R. M. (1991b). Typicality in logically defined categories: Exemplar-similarity versus rule instantiation. *Memory & Cognition*, *19*(2), 131–150. doi:https://doi.org/10.3758/BF03197110

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, *22*(3), 352–369. doi:10.3758/BF03200862

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*(1), 53–79. doi:https://doi.org/10.1037/0033-295X.101.1.53

Redick, T. S., & Lindsey, D. R. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic bulletin & review*, *20*, 1102–1113.

Romberg, A. R., & Saffran, J. R. (2013). All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science*, *37*(7), 1290–1320. doi:https://doi.org/10.1111/cogs.12050

Schlegelmilch, R., & von Helversen, B. (2020). The influence of reward magnitude on stimulus memory and stimulus generalization in categorization decisions. *Journal of Experimental Psychology: General*, *149*(10), 1823–1854. doi:https://doi.org/10.1037/xge0000747

Schlegelmilch, R., Wills, A. J., & von Helversen, B. (2021). A cognitive category-learning model of rule abstraction, attention learning, and contextual modulation. *Psychological Review*, 1211–1248. doi:https://doi.org/10.1037/rev0000321

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. doi:https://doi.org/10.3758/s13423-017-1230-y

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and applied*, *75*(13), 1. doi:https://doi.org/10.1037/h0093825

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. (2015). *Package 'afex'*. Vienna.

Vuong, L. C., Meyer, A. S., & Christiansen, M. H. (2016). Concurrent statistical learning of adjacent and nonadjacent dependencies. *Language Learning*, *66*(1), 8–30. doi:https://doi.org/10.1111/lang.12137

Wilson, B., Spierings, M., Ravignani, A., Mueller, J. L., Mintz, T. H., Wijnen, F., ... Rey, A. (2020). Non-adjacent dependency learning in humans and other animals. *Topics in cognitive science*, *12*(3), 843–858. doi:10.1111/tops.12381

Zeigenfuse, M. D., & Lee, M. D. (2010). A general latent assignment approach for modeling psychological contaminants. *Journal of Mathematical Psychology*, *54*(4), 352–362. doi:https://doi.org/10.1016/j.jmp.2010.04.001