Psychonomic Bulletin & Review (2021) 28:1281–1288
https://doi.org/10.3758/s13423-021-01917-w

BRIEF REPORT

Visual statistical learning is modulated by arbitrary and natural categories

Leeland L. Rogers[1] · Su Hyoun Park[1] · Timothy J. Vickery[1]

Accepted: 11 March 2021 / Published online: 31 March 2021

Abstract

Visual statistical learning (VSL) describes the unintentional extraction of statistical regularities from visual environments across time or space, and is typically studied using novel stimuli (e.g., symbols unfamiliar to participants) and using familiarization procedures that are passive or require only basic vigilance. The natural visual world, however, is rich with a variety of complex visual stimuli, and we experience that world in the presence of goal-driven behavior including overt learning of other kinds. To examine how VSL responds to such contexts, we exposed subjects to statistical contingencies as they learned arbitrary categorical mappings of unfamiliar stimuli (fractals, Experiment 1) or familiar stimuli with preexisting categorical boundaries (faces and scenes, Experiment 2). In a familiarization stage, subjects learned by trial and error the arbitrary mappings between stimuli and one of two responses. Unbeknownst to participants, items were paired such that they always appeared together in the stream. Pairs were equally likely to be of the same or different category. In a pair recognition stage to assess VSL, subjects chose between a target pair and a foil pair. In both experiments, subjects' VSL was shaped by arbitrary categories: same-category pairs were learned better than different-category pairs. Natural categories (Experiment 2) also played a role, with subjects learning same-natural-category pairs at higher rates than different-category pairs, an effect that did not interact with arbitrary mappings. We conclude that learning goals of the observer and preexisting knowledge about the structure of the world play powerful roles in the incidental learning of novel statistical information.

Keywords Visual statistical learning · Categorization · Category learning · Implicit learning · Incidental learning

One way sensory systems cope with the complexity of the visual world is through various forms of learning. Extensive previous work demonstrates that statistical structure (spatiotemporal co-occurrence) is rapidly discovered by human observers in the absence of explicit instruction or intentions to learn such structure, a phenomenon broadly known as statistical learning (for review, see Sherman et al., 2020): Repeated exposure to a predictive relationship between stimuli can lead to subsequent familiarity with that relationship, such as when Stimulus A predicts subsequent Stimulus B, even if the observer is neither informed about nor trying to learn such relationships. A second extensively studied learning phenomenon is categorization (Richler & Palmeri, 2014), which enables variability to be simplified based on broad class distinctions. It stands to reason that categorization and statistical learning may interact in important ways, because objects of statistical learning in natural contexts are often likely to be stimuli about which the learner already has extensive knowledge, such as category membership. Furthermore, statistical learning may sometimes occur in contexts in which competing learning goals (e.g., learning the category of stimuli) are ongoing. Finally, there are many attributes of categorical learning and reasoning that are "statistical" in nature, suggesting a potential overlap with spatiotemporal co-occurrence learning. In this paper, we examine the role of categories and category learning in statistical learning. First, how does explicitly and effortfully learning a category mapping during exposure to statistical regularity impact statistical learning? Second, how do preexisting categorical similarities and distinctions impact statistical learning?

Undirected and incidental regularity extraction was first shown in infants, who were able to extract statistical structure without direction from an artificial auditory stream (Saffran et al., 1996). Adults show the same capabilities (Saffran et al.,

✉ Leeland L. Rogers
Leeland.Rogers@gmail.com

[1] Department of Psychological and Brain Sciences, University of Delaware, 108 Wolf Hall, Newark, DE 19716, USA

1999), including in other sensory modalities, such as vision (i.e., visual statistical learning; VSL). During VSL, observers are capable of incidentally extracting regularities from both temporal (Fiser & Aslin, 2002) and spatial contexts (Fiser & Aslin, 2001). VSL may be constantly active, detecting regularities and storing that knowledge without observer intentions, but some evidence suggests that top-down goals can shape statistically learned representations, at least through the route of selective attention (e.g., Baker et al., 2004; Turk-Browne et al., 2005), with attended pairings learned better than unattended ones. However, how preexisting knowledge and explicit learning goals interact with unintentional statistical learning is largely unknown, especially in contexts wherein the impact of such knowledge and goals does not impose obvious constraints on selective attention.

A separate field of research focuses on learning to form broad classes (reviewed by Ashby & Maddox, 2005; Richler & Palmeri, 2014). Research on category learning is focused on the inference of hidden attributes that link large sets of items to support generalization, whereas VSL is generally related to specific spatiotemporal associative and predictive links. Category membership is usually defined as either a shared set of features that define category boundaries, which can include stimulus attributes, conceptual attributes, or simply a common label, or in the context of distances within similarity spaces (Shepard, 1987), or resemblance of discrete characteristics (Tversky, 1977). Categories can also be abstract or, "ad hoc" (Barsalou, 1983)—for example, "things that need to be donated to charity." The primary link amongst these different forms of category learning and reasoning is the presence of shared features that define categorical membership, whether those features are stimulus based (e.g., color or shape), conceptual (e.g., donation items), or action based (e.g., things that can be kicked).

In contrast to VSL, category learning is commonly studied in contexts where participants are directed to learn, feedback or explicit labels are provided, and learning goals are made explicitly clear. However, it is also possible to learn categories in an unsupervised fashion, in manners that suggest that category learning is itself a "statistical" learning phenomenon, at least in part. For example, even infants notice correlations amongst features (Younger & Cohen, 1986) and are sensitive to statistical attributes such as intraclass variability, which shapes generalization (Quinn et al., 1993). Thus, in our view, the key distinction between VSL and category learning is that the former is usually thought of solely in terms of spatiotemporal regularities, whereas the latter concerns regularities amongst feature attributes.

Both VSL and category learning provide useful accounts of how we may organize and make sense of the world on a moment-to-moment basis. VSL may enable us to bind distinct items together into perceptual objects (Lengyel et al., 2021), support compression in working memory (Brady et al., 2009), and perform predictive processing. Category learning enables us to generalize properties from one instance to another, to achieve invariance with respect to feature variation that does not impact categorical membership, and to track collections bound together by common goals, origins, and hidden properties. While these phenomena are thought of and studied in very different ways, they bear some similarities in terms of sensitivity to statistical information and similar roles in prediction. Additionally, statistical learning in the real world is likely to occur with elements about which the learner already has rich semantic knowledge, including category knowledge. However, statistical learning is typically studied using novel stimuli that do not fall into obvious previously learned categories, or do not have an obvious arrangement in a feature space. This leaves open an important question: How might statistical learning be altered as a function of categorical knowledge?

Vickery et al. (2018) found that task demands strongly shaped what is statistically learned, and that category information can influence VSL when attended. In their experiment, participants viewed a stream of face and scene images that covertly appeared in pairs. Participants received one of two sets of instruction: One group was told to respond when an image "jiggles" back and forth while the other was told to categorize the image with one of two button presses (e.g., female faces and an outdoor image could share one response while male faces and indoor scenes shared the other response). In the detection group, no significant differences in learning were observed between different stimulus combinations. However, in the categorization group, when two stimuli of a pair shared a response, greater learning was observed than when they did not, even when the items did not share category. Learning was also better for same-category than for different-category pairs, at least if the items shared the same response. This experiment suggested that natural category membership and/or response mapping had powerful impacts on VSL. However, two important limitations of the study limited our conclusions. First, response mapping in categorization was confounded, at least within same-category pairs, with category membership and similarity. That is, it is likely that female faces are more like one another than they are to male faces, and that indoor scenes are more like other indoor scenes than outdoor scenes, on average. Secondly, while no distinctions among category combinations were observed within the "jiggle" detection group, this could have been due to reduced depth of processing and/or a floor effect that limited variability in learning across combination types. That is, detecting a jiggle required rather shallow processing relative to categorization, which requires processing of features to match to a subcategory. This alone, rather than the specific act of categorizing the items, might have induced an effect of categorization on learning.

In the current paper, we isolated the role of response mapping from visual similarity by arbitrarily mapping diverse stimuli onto two keys during a "category" learning stage that also served as a VSL familiarization stage. To limit the impact of all factors other than category assignment, categories were constructed in a spare fashion—by simply linking one of two labels/responses arbitrarily to each item during VSL familiarization of paired items. Previewing our results, we found that, in two experiments, arbitrary category assignments strongly shaped what pairs were remembered—same-category pairs were correctly chosen at higher rates than different-category pairs during a test phase. Experiment 2 further demonstrated that natural categories shape memories for pairings in the same manner, and that effect did not interact with arbitrary category mappings.

Experiment 1

In Experiment 1, we employed fractal images that had no obvious or systematic interrelationships in feature space. During familiarization, subjects learned to map those images onto one of two keys by trial-and-error learning. Subjects were uninformed that stimuli also appeared in pairs, such that Fractal A always preceded B. Such pairs were evenly divided between same-response and different-response mappings, allowing us to examine whether response mapping determined what statistical relationships were learned (assessed in a surprise recognition memory phase), in the absence of any clear relationship between visual similarity and arbitrary categories defined by key mappings.

Methods

Participants

Sample sizes were based on power analysis derived from Vickery et al. (2018). We based our analysis on Experiment 4 of that paper, specifically the categorization subgroup (whose familiarization task resembled the one used here) and the effect of response (averaging over "task"). Cohen's d_z was 0.692 for this comparison. To obtain a power of .95 to detect an effect of this magnitude with a two-tailed *t* test would require a sample size of 30 subjects, so this was selected as our target sample for both experiments of the current study.

Our focus being on subsequent memory for pairs as a function of category learning, participants who failed to reach an overall mean of 60% accuracy during the first, training phase were excluded from analysis. Participants were expected to easily reach well-above-chance levels of learning, so this cutoff was arbitrarily selected to act as a liberal criterion for participant inclusion. No participants were excluded from Experiment 1, and two participants were excluded from Experiment 2 based on this criterion. Excluded participants were replaced to achieve the target sample size. Across the two experiments, a total of 62 University of Delaware undergraduate students (ages 18–22 years) were recruited. Experiment 1 (N = 30) included 25 female participants, four male participants, and one participant whose gender was not disclosed. Experiment 2 (usable N = 30, after excluding two subjects) were not asked to disclose gender but were sampled from the same participant pool as Experiment 1. Subjects who completed Experiment 1 were precluded from completing Experiment 2. All experiments were conducted with the informed consent of subjects and with approval of the Institutional Review Board at University of Delaware.

Stimuli and materials

Stimuli consisted of 32 fractal images, which were gathered in-house from various web sites, before the study. The images were randomly and covertly assigned to 16 pairs independently for each subject. Images were further randomly assigned for each subject to two equal-sized category or key mapping groups (z images associated with the 'z' key and m images associated with the 'm' key), with the constraint that eight pairs were same-category and eight pairs were different-category (we refer to these as "pair types"), with equal representation of each category in each position in the pair. An example subset of these assignments and same/different pairings for one hypothetical subject are displayed in Fig. 1a–b. The use of unfamiliar and diverse fractal images with no natural categorical membership, and the use of random assignment, ensured that visual similarity was equal across different pair types and within and between categories, on average. All images were presented on a 24-inch diagonal LED monitor (120 Hz) with a resolution of 1,920 pixels × 1,080 pixels from an unrestrained viewing distance of approximately 57 cm. Images were presented at 200 pixels × 200 pixels (approximately 5.5° × 5.5°).

Procedure

After providing informed consent, participants were seated at a computer. They then completed two phases (familiarization and a surprise test phase), with instructions presented before each phase.

Familiarization phase During this phase, images were presented one at a time, and subjects were instructed to learn, by trial and error, the category (z or m) to which the image belonged, by pressing the 'z' or 'm' key with their left or right index finger, respectively. Each image was presented onscreen for 1 second and participants were required to respond before the image disappeared. Feedback was provided in the form of a green fixation circle for correct responses that remained onscreen for 1 second. If incorrect, participants were presented with a red fixation circle for 1.5 seconds (the additional time

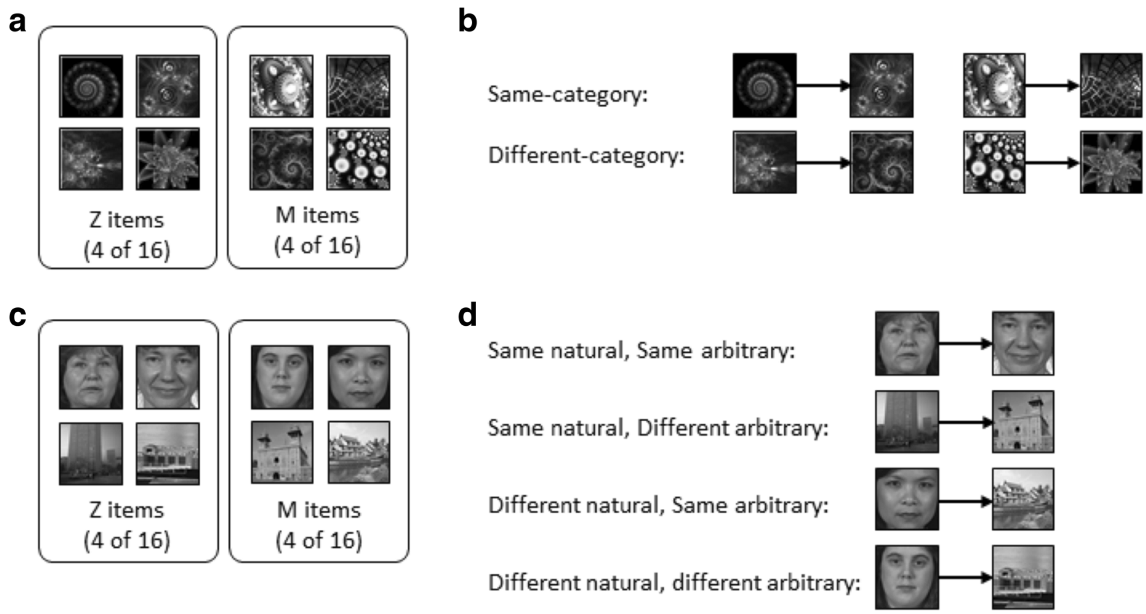

Fig. 1 Experiments 1 and 2 stimuli pairing and assignment to categories. *Note.* Hypothetical examples of random stimulus assignments to categories and pairings. **a** Experiment 1, four of 16 images assigned to each category. **b** Experiment 1, from items in a, four different pairs used to construct sequences exposed during familiarization. The arrow indicates the sequence in which the images within a pair would occur during familiarization. There were eight total same-category and eight different-category pairings. There were eight pairs in each of the two pairing types. **c** Experiment 2, four of 16 images assigned to each category. **d** Experiment 2, from items in **c**, four different pairings used to construct sequences exposed during familiarization. The arrow indicates the sequence in which the images within a pair would occur. There were four total pairs for each of the four types of pairings

served as an incentive to learn). Sequences were constrained such that all images appeared in their respective pair orders (e.g., Image A always immediately preceded Image B in Pairing AB), and pair orders were pseudorandomized such that no pair could immediately repeat or repeat with a single intervening pair. Participants viewed each pair four times per block across six total blocks of training. All images appeared an equal number of times (24 in total). Subjects were given untimed breaks between blocks.

Test phase Following completion of the familiarization phase, subjects were informed that the images encountered during that phase always appeared in pairs. On each trial of a surprise test phase, subjects viewed two sequences of pairs of images, and were instructed to choose the pair that appeared more familiar to them by pressing the '1' or '2' key. One pair was a target pair that had appeared during familiarization multiple times, and the second pair was a foil pair composed of two images that did not appear together in sequence previously. The first image of a foil pair was taken from the set of first images of target pairs, while the second image of a foil pair was taken from the set of second images of different target pairs. Foil pairs (16) were fixed and appeared an equal number of times as each target pair during test, to prevent further pair learning during the test phase. All images appeared an equal number of times at a rate equal to that experienced during familiarization (1 second on, 1 second off), along with indicators that a sequence was beginning (e.g., "Sequence 1," 1 second on, 1 second off) that preceded each pair. Target pairs were equally likely to appear first or second in the sequence. The test phase consisted of 64 forced-choice trials that were unspeeded, and subjects did not receive feedback during the test phase.

Results

Familiarization phase

Categorization performance during familiarization is shown in Fig. 2a. Participants reached above-chance accuracy within the first block, $t(29) = 2.08$, $p = .046$, $d = 0.38$ (one-sample t test of proportion correct vs. chance level of 0.5) and remained above chance through the sixth block (all ps < .05).

A one-way repeated-measures ANOVA, with accuracy as the dependent variable (DV) and block (six levels) as the IV, was conducted to examine whether category learning improved over time. There was a significant effect of block, $F(5, 145) = 191$, $p < 0.001$, $\eta_p^2 = 0.868$. Post hoc t tests comparing all blocks' performance with p values corrected using Holm's method indicated that all blocks' performance differed from all other blocks (all ps < .05), except for Blocks 5 and 6 ($p = .22$). Thus, performance improved steadily until it leveled off by around Block 5, with performance numerically peaking at Block 6 (proportion correct, $M = 0.90$, $SD = 0.090$). A detailed analysis of accuracy and response time (RT) during training, separated by position within a sequence (first and

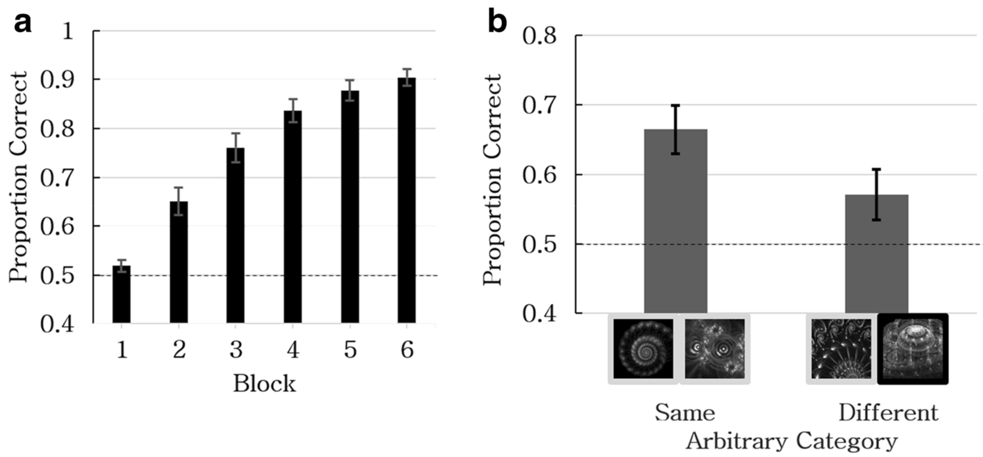

Fig. 2 Familiarization proportion correct and test proportion correct data for Experiment 1. *Note.* **a** Mean categorization accuracy during the familiarization phase from Experiment 1. **b** Mean proportion correct from Experiment 1 for same-group pairs and different-group pairs. The outlines reflect hypothetical arbitrary category assignment. Participants did not see outlines. Error bars depict the standard error of the mean

second) and pair type (same or different), is presented in the Supplementary Materials. In short, we found no effect of order or interaction with order on accuracy in the last two blocks, though a more complex interaction (likely due to subjects' learning strategies) was present in early blocks. In terms of RT, subjects were slightly faster to respond to second items in a sequence versus first items ($p = .043$), but sequence position did not significantly interact with pair type.

Test phase

Mean accuracy at identifying the old pair was computed for each subject and target pair type during the test phase (see Fig. 2b). These values were transformed using the arcsine-square-root transformation, and statistically compared with chance performance (accuracy of 0.5) using a one-sample t test for each type, and across pair type (same vs. different category) using a paired t test. Both same-category, $t(29) = 4.99$, $p < .001$, Cohen's $d = 0.912$, and different-category, $t(29) = 2.19$, $p = .037$, $d = 0.4$, pair types were chosen at above-chance rates. There was a significant difference between these two rates of recognition, however, $t(29) = 2.76$, $p = .01$, Cohen's $d_z = 0.504$. Same-category pairs were recognized at higher rates ($M = .621$, $SD = .130$) than different-category pairs ($M = .550$, $SD = .125$). The Supplementary Materials presents a corresponding mixed-effect logistic regression analysis of this data, which yields the same conclusions.

Discussion

Subjects were adept at learning arbitrary category mappings of images. They further demonstrated learning of pairings for both same-category and different-category pairs. However, learned arbitrary categories predicted the strength of VSL such that same arbitrary category pairings were learned better than different-category pairings. Thus, newly learned arbitrary information seems to have enhanced VSL when two items of a pair shared the same category. This difference cannot be attributed to prior knowledge of the stimuli (before training), as all images are novel to participants and did not consistently fall into any kind of naturally occurring category.

A number of potential mechanisms may be responsible for this result. One possibility is that the simple fact that two items share a label may make items more similar in their memory representations, which may in turn lead to greater tendency for those items to form associative links in memory. Another possibility is that the act of changing category responses across first and second items in different-category pairs could disrupt working memory storage of the first item, such that the first and second items are less associated because their common storage in working memory was disrupted. Similarly, the change in category from first to second items for different-category but not same-category pairs could induce an "event boundary." Event boundaries are known to have effects on long-term memory (DuBrow & Davachi, 2016).

An obvious question arises from these results: Does prior categorical knowledge shape VSL like the novel arbitrary information learned in Experiment 1? Would such preexisting knowledge and experience interact with the arbitrary categorical information that is more recently imposed on the stimuli?

Experiment 2

With Experiment 2, we aimed to replicate the finding of Experiment 1, and further examine how natural categories impact VSL, when the task during learning requires processing of images to individuate and identify them, but is orthogonal to natural category assignment. To this end, we maintained our basic design from Experiment 1, but employed face and scene images in place of the fractal images. Thus, face and scene served as an orthogonal natural category membership factor, which we manipulated and crossed with arbitrary category membership.

Methods

Methods were identical to those of Experiment 1, except for the stimuli used, which consisted of color images of faces and scenes, and the way these images were divided into pairs was affected by this. Here, we only highlight method differences.

Stimuli and materials

Stimuli were 16 face from the FERET database (Phillips et al., 1998) and 16 scene images from Vickery et al. (2018) that were originally collected from the internet. These images were of the same size as fractals in Experiment 1 and were randomly sampled for each participant from a pool of 50 images of each type. Participants were randomly assigned either male or female images, and either indoor or outdoor scenes. The set of images selected for the participant were divided into 16 pairs with the following constraints: eight pairs were same-arbitrary and were assigned to the same key during familiarization (categorization training), while eight pairs were different-arbitrary, assigned to different keys. Of the eight pairs of each arbitrary type, four were same-natural-category (two both-faces and two both-scenes), and four were different-natural-category (two were face followed by a scene and two pairs were scene followed by a face). Thus, there were two crossed, orthogonal factors associated with each pair: arbitrary category (same or different) and natural category (same or different). Assignment to key was counterbalanced across all four pair types defined by these factors. An example subset of these assignments and same/different pairings are displayed in Fig. 1c–d.

Results

Familiarization phase

A one-way repeated-measures ANOVA, with block as a factor, was used to examine category leaning during familiarization (see Fig. 3a). Categorization accuracy changed across blocks, $F(5, 145) = 211$, $p < .001$, $\eta_p^2 = 0.879$, with participants reaching above-chance accuracy by the second block, $t(29) = 7.29$, $p = .001$, $d = 1.33$. Post hoc tests comparing all blocks demonstrated that all blocks were significantly different from one another in terms of categorization performance (all ps < .006, corrected for multiple comparisons using Holm's method)—performance increased monotonically from block to block.

A more detailed analysis of accuracy and RT during familiarization is presented in the Supplementary Materials. In short, we focused on the final two blocks, and found no effect of position on categorization accuracy, as well as no interaction of position with other factors on categorization accuracy. We did, however, find an effect of position on RT ($p < .001$), with faster responses to second items than to first items in a sequence, consistent with a predictive response. Position effects interacted with natural category composition, however ($p = .009$), such that order effects only manifested for different-natural-category pairs and not same-natural-category pairs. More details and discussion are presented in the Supplementary Materials.

Test phase

A 2 × 2 RM-ANOVA was conducted on arcsine-square-root transformed accuracy during the test phase, with factors of arbitrary pair type (same-arbitrary vs. different-arbitrary) and natural category pair type (same-natural vs. different-natural). The four available pair-type conditions were same-natural same-arbitrary ($M = 0.740$, $SD = 0.154$), same-natural different-arbitrary ($M = 0.662$, $SD = 0.182$), different-natural same-arbitrary ($M = 0.648$, $SD = 0.175$), and different-natural, different-arbitrary ($M = 0.568$, $SD = 0.189$). Significant main effects of arbitrary category pair type, $F(1, 29) = 15.23$, $p < .001$, $\eta_p^2 = .344$, and natural category pair type, $F(1, 29) = 6.07$, $p = .02$, $\eta_p^2 = .173$, were observed. In both cases, same pairs were recognized at higher rates than different pairs. However, the interaction was not significant, $F(1, 29) = 0.07$, $p = .793$, $\eta_p^2 = .002$. The Supplementary Materials presents a corresponding mixed-effect logistic regression analysis of this data, which yields the same conclusions.

Discussion

As in Experiment 1, we observed greater learning for pairs that shared an arbitrary category as defined during familiarization. Additionally, we observed greater learning for pairs that shared a natural category, despite natural category distinctions being irrelevant to the task at hand. There was no interaction between influences from arbitrary categories and natural category membership, suggesting that the total information shared between two images of a pair predicts the ability to statistically learn the pair. Natural category membership shapes learning even when it is unattended, at least if the cover task requires individuation and identification of items during familiarization.

This result suggests that the act of changing category response, offered as one explanation for Experiment 1's results, cannot solely explain the same versus different category difference because a change in natural categories did not result in

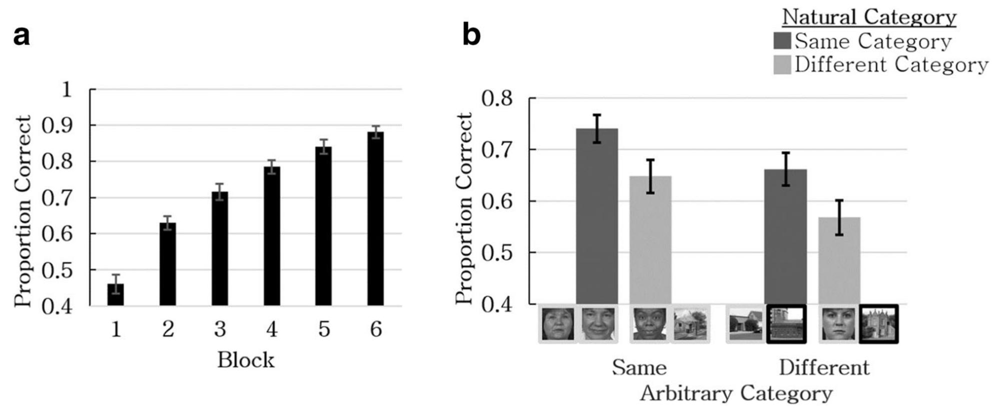

Fig. 3 Familiarization proportion correct and test proportion correct data for Experiment 2. *Note.* **a** Mean categorization accuracy during the familiarization phase from Experiment 1. **b** Mean proportion correct from Experiment 2 for same/different arbitrary group pairs and same/different natural category pairs. Outlines around images indicate their arbitrary group membership. Participants did not see these outlines. Error bars depict the standard error of the mean

a category response change. In our view, this makes it more probable that similarity of representations (shaped by both novel and preexisting category knowledge) in memory accounts for most of the effect of categories. However, it is still possible that the natural category change across different-category pairs induced an event boundary that shaped associative pair memory, even though such distinctions were not cued explicitly by the task.

General discussion

These experiments highlight the importance of category learning and prior knowledge in VSL by demonstrating the impact of recently learned groupings and the influence of task-irrelevant and familiar natural category information. When items in a pair came from the same category, whether arbitrary or natural, they were learned better than when they came from different categories. This was true whether the categorical distinction was new to observers and learned during the same familiarization period that exposed subjects to statistical information, or whether it was a previously learned distinction that had no bearing on the familiarization task.

Vickery et al. (2018) found that category membership impacted performance. However, it was unclear whether this was due to visual similarity confounds with category membership. In Experiment 1, we randomly assigned fractals to two arbitrary categories and asked subjects to learn mappings of those fractals while they were exposed to statistical regularities. Because fractals were randomly assigned to category, within-category similarity and between-category similarity were balanced on average. The fact that these arbitrary category mappings had a powerful impact on learning demonstrates that similarity is not likely the sole factor contributing to prior results.

Another unresolved question from Vickery et al. (2018) work was whether category information had to be explicitly attended to impact statistical learning performance. Apparently, this was not the case—two-face or two-scene pairs were learned better than face and scene pairs in Experiment 2. This contrasts with Vickery and colleagues' finding, that when same-category and different-category pairs were exposed during a familiarization phase, with a cover task to detect an image "jiggle," no differences were observed in learning of those types of pairings. One possible reason for this is floor effects—learning was generally weak across pair types under the detection task in this prior work, and it is possible that with increased exposure and better overall learning these patterns might emerge even under a detection cover task. Another possibility is that the individuation and identification of images, which was required by our categorization cover task, is important to induce category effects.

It is important to mention that Experiment 2 did not separate natural category information from visual similarity. The current work cannot disentangle the role of natural category membership from that of visual similarity (i.e., a pair consisting of two faces share a great deal of visual information above and beyond their category as faces). Experiment 1 suggests that category information alone can induce differences in learning, though, in the absence of similarity differences. Future work will be needed to dissociate visual similarity information from categorical information, to assess the true impact of item similarity on VSL.

These findings may have profound implications for how real-world statistical learning occurs. We have shown that both recently learned and well-established category distinctions can predict the strength of statistically learned representations. This suggests that prior knowledge may have a powerful impact on statistical learning, and the strength of that prior knowledge may exert influences on learning despite being irrelevant to the task at hand. Our results also add to growing evidence that the task during exposure to statistical regularities plays an important role (Turk-Browne et al., 2005; Vickery et al., 2018; Zhao et al., 2011; Zhao et al., 2013; Zhao & Yu, 2016).

In the natural world, it is highly likely that most statistical regularities are encountered under contexts in which those regularities have no obvious bearing on the immediate task at hand (e.g., even in navigation tasks, we may follow a series of landmarks described in directions while encountering numerous other statistical regularities that are not relevant to navigation, per se). In addition, statistical learning must often occur over stimuli that are associated with deep prior knowledge and experience. Thus, understanding how tasks and prior knowledge impact statistical learning is essential to developing an appreciation for how it may manifest in real-world contexts.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13423-021-01917-w.

Author note This research was supported by National Sciences Foundation grants BCS 1558535 and OIA 1632849 to T.J.V.

Data availability The data, but not the related materials, for all experiments are available (https://osf.io/hf7tv). None of the experiments were preregistered.

References

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*(1), 149–178. https://doi.org/10.1146/annurev.psych.56.091103.070217

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Visual statistical learning role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, *15*(7), 1–8.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*(3), 211–227. https://doi.org/10.3758/BF03196968

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology. General*, *138*(4), 487–502. https://doi.org/10.1167/8.6.199

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory, 134*, 107–114. https://doi.org/10.1016/j.nlm.2016.07.011

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499–504. https://doi.org/10.1111/1467-9280.00392

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458–467. https://doi.org/10.1037/0278-7393.28.3.458

Lengyel, G., Nagy, M., & Fiser, J. (2021). Statistically defined visual chunks engage object-based attention. *Nature Communications*, *12*(1), 1–12. https://doi.org/10.1038/s41467-020-20589-z

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, *16*(5), 295–306.

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, *22*(4), 463–475. https://doi.org/10.1068/p220463

Richler, J. J., & Palmeri, T. J. (2014). Visual category learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(1), 75–94. https://doi.org/10.1002/wcs.1268

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317–1323. https://doi.org/10.1126/science.3629243

Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, *32*, 15–20. https://doi.org/10.1016/j.cobeha.2020.01.015

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327–352. https://doi.org/10.1037/0033-295X.84.4.327

Vickery, T. J., Park, S. H., Gupta, J., & Berryhill, M. E. (2018). Tasks determine what is learned in visual statistical learning. *Psychonomic Bulletin & Review*, *25*(5), 1847–1854. https://doi.org/10.3758/s13423-017-1405-6

Younger, B. A., & Cohen, L. B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development, 57*(3), 803–815. https://doi.org/10.2307/1130356

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological Science*, *24*(5), 667–677. https://doi.org/10.1177/0956797612460407

Zhao, J. Ngo, N., McKendrick, R., & Turk-Browne, N. B. (2011). Mutual interference between statistical summary perception and statistical learning. *Psychological Science*, *22*(9), 1212–1219. https://doi.org/10.1177/0956797611419304

Zhao, J. & Yu, R. Q. (2016). Statistical regularities reduce perceived numerosity. *Cognition*, *146*, 217–222. https://doi.org/10.1016/j.cognition.2015.09.018