

Unitization, similarity, and overt attention in categorization and exposure

Alice K. Welham · Andy J. Wills

© Psychonomic Society, Inc. 2011

Abstract Unitization, the creation of new stimulus features by the fusion of preexisting features, is one of the hypothesized processes of perceptual learning (Goldstone Annual Review of Psychology, 49:585–612, 1998). Some argue that unitization occurs to the extent that it is required for successful task performance (e.g., Shiffrin & Lightfoot, 1997), while others argue that unitization is largely independent of functionality (e.g., McLaren & Mackintosh Animal Learning & Behavior, 30:177–200, 2000). Across three experiments, employing supervised category learning and unsupervised exposure, we investigated three predictions of the McLaren and Mackintosh (Animal Learning & Behavior, 30:177–200, 2000) model: (1) Unitization is accompanied by an initial increase in the subjective similarity of stimuli sharing a unitized component; (2) unitization of a configuration occurs through exposure to its components, even when the task does not require it; (3) as unitization approaches completion, salience of the unitized component may be reduced. Our data supported these

predictions. We also found that unitization is associated with increases in overt attention to the unitized component, as measured through eye tracking.

Keywords Categorization · Associative learning · Perception

Classic models of categorization tend to assume that objects are represented by a fixed set of perceptual components. Whether those components are discrete features (e.g., Medin & Schaffer, 1978) or continuous dimensions (e.g., Nosofsky, 1986), the process of categorization does not add to or subtract from the fixed set of perceptual components in such models, although it is generally accepted that categorization is able to affect the amount of attention these preexisting dimensions or features attract. This *fixed feature* approach to categorization has been contrasted with a *feature creation* approach, which hypothesizes that the features one perceives in stimuli can qualitatively change as a result of experience with those stimuli (Schyns, Goldstone, & Thibaut, 1998).

In the present article, we investigate one of the ways in which new features might develop during categorization: the fusion of a set of preexisting features. This fusion is often described as *unitization* (e.g., Goldstone, 1998; McLaren & Mackintosh, 2000) and has been hypothesized to underlie the word superiority effect (Cattell, 1886), the object superiority effect (Weisstein & Harris, 1974), and the effects of extended practice on conjunctive visual search (Shiffrin & Lightfoot, 1997).

One of the few investigations of unitization occurring as a consequence of category learning was presented by Goldstone (2000). In his experiments, participants learned to categorize stimuli that consisted of contiguous segments of curves. In one condition, the categorization could be

This research was supported by BBSRC Grant 9/S17109 and an ESRC studentship grant. The authors thank Rob Goldstone, Ian McLaren, Stephen Monsell, and Richard Shiffrin for their helpful comments on this work.

Electronic supplementary material The online version of this article (doi:10.3758/s13421-011-0124-x) contains supplementary material, which is available to authorized users.

A. K. Welham · A. J. Wills (✉)
University of Exeter,
Perry Road,
Exeter EX4 4QG, UK
e-mail: a.j.wills@ex.ac.uk

Present Address:
A. K. Welham
School of Psychology, University of Birmingham,
Birmingham, UK

performed on the basis of just one segment of curve; in the other condition, the categorization could be performed only on the basis of all five curve segments. Reductions in reaction time were found to be more substantial for the all-segments condition than for the one-segment condition, an effect that occurred only when the components of the conjunction were presented in a fixed configuration (as opposed to being randomly reordered on each trial). Employing a control task in which component detection was not required, and through the application of Fourier deconvolution analysis of reaction times (a variant of the Donders subtractive method; Donders, 1969/1868–1869; see Sheu & Ratcliff, 1995, for a critique), Goldstone (2000) concluded that the improvement in reaction time in the all-segments condition was larger than would be expected if the stimulus components were being processed independently. He therefore concluded that unitization had occurred in the all-segments conditions.

Predictions of the McLaren–Mackintosh theory of unitization

If one accepts Goldstone's (2000) analysis, the implication is that exposure to stimulus components during categorization can lead to those components being processed more quickly than one would predict if each component were being processed independently. One explanation of how this might occur is that the detection of one component facilitates detection of the other components with which it reliably co-occurs (Goldstone, 2000, p. 110). A number of simple connectionist models exist that could account for the development of facilitation of this kind (e.g., McClelland & Rumelhart, 1985). According to these models, repeated presentation of a componential stimulus will cause associative links to form between those components. Once such connections have been formed, the detection of one stimulus component should facilitate detection of the others through spreading activation. The idea that such within-compound associations do indeed form has been supported by a number of experiments in which spatially discrete components reliably co-occur and knowledge of the co-occurrences is tested either directly (Melchers, Lachnit, & Shanks, 2004) or indirectly through, for example, familiarity judgments of seen and unseen conjunctions (Fiser & Aislin, 2001; Larkin, Aitken, & Dickinson, 1998; Wasserman & Berglan, 1998).

One particular connectionist model, that of McLaren and Mackintosh (2000, 2002), was of particular interest in the present investigation because it had been designed specifically as an account of perceptual learning, and unitization is often considered to be one of the processes by which perceptual learning occurs (Goldstone, 1998). Support for

the McLaren–Mackintosh model as an account of human perceptual learning comes from a number of its correct but nonintuitive predictions. For example, it correctly predicts that whether stimulus preexposure facilitates or retards subsequent categorization depends on the way the category examples are created from the base patterns (Wills & McLaren, 1998; Wills, Suret, & McLaren, 2004). The McLaren–Mackintosh model also makes a number of as-yet-untested predictions about the necessary conditions for, and the consequences of, unitization, and it is these predictions that are investigated in the present article.

The McLaren–Mackintosh model takes as its starting point the parallel distributed processing approach of McClelland and Rumelhart (1985). Each stimulus is represented as a pattern of activation across a set of input units that represent the elements of the stimulus. These elements are assumed to represent small segments of stimuli, and even simple stimuli are represented across multiple units. Following stimulus-sampling theory (Estes, 1950), a central assumption of the McLaren–Mackintosh model is that only a subset of the elements of a stimulus are sampled on any given trial. The more complex the stimulus, the smaller the proportion of elements sampled, and the higher the variability of that sample between trials (McLaren & Mackintosh, 2000, p. 212).

Unitization occurs in the McLaren–Mackintosh model in the same manner as it does in a number of other connectionist theories (e.g., McClelland & Rumelhart, 1985). Specifically, assume that units i and j represent different elements of stimulus X and that i and j have not appeared in conjunction before. The McLaren–Mackintosh model is such that when these units are reliably externally coactivated, the connections between them strengthen. Over time, a network of connections will develop between elements that co-occur, and this is the sense in which stimuli are unitized in the McLaren–Mackintosh model.

One consequence of unitization in the McLaren–Mackintosh model is that unitized stimuli activate a greater proportion of their representational elements than do nonunitized stimuli. As was stated above, the presentation of a stimulus is assumed to directly activate only a subset of its representational elements on any given trial. However, the associative connections formed between representational elements as a result of unitization mean that the directly activated elements can indirectly activate further representational elements of the stimulus (through the associative links that have formed between those elements). This increase in the proportion of representational elements activated as a result of unitization is predicted by the McLaren–Mackintosh theory to affect the subjective similarity of stimuli. Below, we illustrate this prediction through an example.

As an elemental theory, similarity in McLaren–Mackintosh is assumed to be an increasing function of the number of elements two stimuli have in common (cf. Tversky, 1977). Consider two stimuli that share a component (a collection of representational elements) that reliably recurs from trial to trial but that, otherwise, consist of trial-unique noise; such stimuli are illustrated in Fig. 1. For the purposes of this example, assume that the stimuli shown in Fig. 1 are represented by a set of 256 elements, 1 for each of the squares of the stimulus. Assume that the reliably recurring component makes up 25% of the representational elements and that 10% of the representation elements are directly activated on any given trial (25% and 10% are used for illustration; any values greater than 0% and less than 100% will suffice).

Prior to unitization, two stimuli sharing the reliably recurring component but otherwise made up of noise will, on average, have 62.5% of their active representational elements in common (all of the reliably co-occurring elements sampled plus, by chance, half of the noise elements sampled). After complete unitization, the same stimuli share 88.5% of the representational elements. This is because the directly sampled 10% of elements of the reliably recurring component associatively activate the remaining 90% of elements of the reliably recurring component. Hence, the McLaren–Mackintosh theory predicts that unitization will result in an increase in the subjective similarity of two stimuli that share a reliably recurring component. Experiment 1 investigated this prediction through a categorization task in which these reliably recurring components perfectly predicted category membership. The effects of this categorization training on the rated similarity of pairs of stimuli sharing a reliably recurring component are then examined. Under the McLaren–Mackintosh theory of unitization, the rated similarity should rise as a result of the categorization training.

In addition to the unitization mechanism, the McLaren–Mackintosh model also has a salience reduction mecha-

nism. The salience reduction mechanism reduces the activation level of stimulus elements to the extent that they are well predicted by other stimulus elements. This aspect of the model is critical to its ability to predict the finding that stimulus structure modulates the effect of preexposure, with some stimulus structures leading to enhanced categorization, and others leading to retarded categorization (Wills & McLaren, 1998; Wills et al., 2004). The elements of a unitized stimulus are, of course, well predicted by the other elements of that stimulus, so there are mechanisms in the model that would tend to reduce the activation level of unitized stimulus elements. This reduction in activation level might, in turn, lead to a reduction in the subjective similarity of stimuli sharing a unitized component (because the elements that they share will have a lowered activity level and, thus, a smaller contribution to similarity as indexed by shared elements). The combination of the unitization process and the salience reduction process means that the McLaren–Mackintosh model predicts exposure leads to an *initial* increase in the similarity of nonunitized configuration-sharing pairs (through unitization, which results in an increase in the number of active representational elements), but the model also predicts that, as unitization approaches completion, the similarity of configuration-sharing pairs should be reduced (due to the reduction in the activity level of the unitized elements via the salience reduction process). In order to examine this prediction, all the experiments in the present study included a condition where the configurations were obvious from the outset (see Fig. 2 for examples). Our hypothesis is that these configurations unitize quickly and, hence, exposure to these configurations should reveal mainly a salience reduction process (and therefore, the similarity of configuration-sharing pairs should be reduced). We further hypothesize that the nonobvious configurations (Fig. 1) will take longer to unitize and, hence, exposure to nonobvious configurations should reveal mainly a unitization process (and therefore, the similarity of configuration-sharing pairs should increase).

Among theories of unitization, the McLaren–Mackintosh account is unique in that it predicts that unitization of a repeatedly presented set of components should occur whether or not the set of components is diagnostic of some category. Despite longstanding evidence that perceptual



Fig. 1 Examples of the stimuli used in the nonobvious configuration conditions of Experiments 1–3. Each training stimulus contained one of two category-defining configurations, in one of four locations. The rest of the stimulus comprised trial-unique noise. For the purposes of illustration, the nonobvious configuration is highlighted in each of the Category I examples in this figure. Configurations were not highlighted in the stimuli actually presented to participants

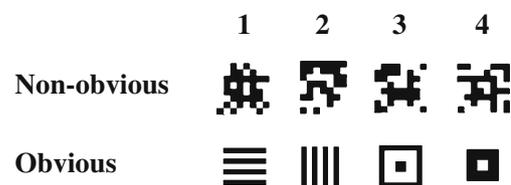


Fig. 2 The non-obvious and obvious configurations employed in Experiments 1–3

change can occur as a result of simple exposure (e.g., Gibson & Walk, 1956), there has been an emphasis in recent years on the role of functionality in feature creation. For example, Schyns, Goldstone, & Thibaut (1998) argued that “features should be functionally determined” (p. 4). Goldstone (2000) stated that his experiments investigated “whether new perceptual units can be developed if they are useful for a category learning task” (p. 86), and Shiffrin and Lightfoot (1997) suggested that “unitization may not occur unless the task requires unitization for success” (p. 74). In contrast, models such as McLaren–Mackintosh assume that unitization of features occurs through simple exposure, independently of any function those features may have. In the McLaren–Mackintosh account, it is the stimulus exposure in categorization that causes unitization to occur, and therefore, unitization should also be observed where stimuli are exposed with no requirement to categorize. This prediction was tested in Experiment 2. Experiment 2 followed the same basic procedure as that in Experiment 1 but replaced the categorization training with an exposure task (participants were simply asked to rate the “attractiveness” of the presented stimuli, rather than to categorize them).

Experiment 2 also served a second, related function. In assessing the changes in perceived similarity that may be brought about by unitization, it is important to dissociate the hypothesized unitization-related changes from changes that might be expected to occur for other reasons. In particular, the amount of overt attention attracted by even very familiar stimuli (e.g., words) is sometimes modulated by the extent to which those stimuli predict an outcome (e.g., Kruschke, Kappeman, & Hetrick, 2005; Rehder & Hoffman, 2005). Specifically, those stimuli that reliably predict an outcome sometimes attract more overt attention than do those that do not. In the categorization of stimuli that comprise a reliably recurring component plus trial-unique noise, the recurring component may come to attract more attention than the rest of the stimulus, as a consequence of it being the only part of the stimulus that reliably predicts the category label. It is not unreasonable to assume that an increase in attention to the recurring component would increase the rated similarity of two stimuli sharing that component. By exposing stimuli with no requirement to categorize them, Experiment 2 acted as a control for any effects of outcome-prediction-based increases in overt attention that may have occurred in Experiment 1.

A related issue is the possible presence of *categorical perception* effects in Experiment 1—in other words, the finding that the act of categorizing stimuli can lead to between-category differences being accentuated and within-category differences being deemphasized (see Goldstone & Hendrickson, 2010, for a review). Multiple theories of the

process(es) underlying categorical perception exist, and predictability-based changes in overt attention (discussed above) are one possible process in some cases. The central point for the present experiments was that one empirical effect of the broad phenomenon of categorical perception might plausibly be to increase the similarity of two stimuli that are considered to belong to the same category. By exposing stimuli with no requirement to categorize them, Experiment 2 acted as a control for the effects of categorical perception.

Our final experiment, Experiment 3, served two purposes. First, it acted as a partial within-experiment replication of Experiments 1 and 2. Second, it investigated predictions of the unitization and salience reduction processes of the McLaren–Mackintosh theory, using an additional methodology: the measurement of overt attention as indexed by eye gaze. Specifically, it examined changes in overt attention to the repeated configuration during exposure, as a function of configuration type (obvious vs. nonobvious; Fig. 2) and exposure type (categorization, as in Experiment 1, vs. exposure, as in Experiment 2). Exposure to nonobvious configurations was expected to increase overt attention to the configuration (through the unitization process), while exposure to obvious configurations was expected to decrease overt attention to the configuration (through the salience reduction process).

Experiment 1

Our first experiment explored the consequences of unitization on postcategorization similarity. The experiment comprised a training phase, in which participants were taught to categorize stimuli, followed by a test phase in which similarity judgments were collected. Every participant was exposed to two configurations in the training phase, and each configuration was fully predictive of one category. In the test phase, similarity judgments were collected. Of principal interest were any changes in the similarity of pairs of stimuli sharing trained configurations, as compared with pairs of stimuli sharing nontrained configurations. Two assays of the similarity of these *configuration-sharing pairs* were employed: a numerical rating of the similarity of sequentially presented stimuli and a forced choice measure of similarity of simultaneously presented stimuli. We used more than one measure to provide the possibility of converging evidence within this relatively novel experimental paradigm; we had no particular reason to expect qualitatively different patterns of results from these two tests.

The test phase also included an examination of *noise-sharing* stimulus pairs. Noise-sharing stimulus pairs differed only in terms of the configuration: One stimulus of

the pair contained a configuration, and the other did not. The thinking behind these noise-sharing pairs was that, under the McLaren–Mackintosh theory, unitization might be expected to result in a reduction of the similarity of these stimulus pairs, because the increased number of representational elements activated by the configuration (as a result of unitization) would decrease the proportion of representational elements shared by the two stimuli.

As is illustrated in Fig. 1, the stimuli we employed were 16×16 black-and-white checkerboards. This stimulus type was chosen for its perceptual complexity (relative to stimuli often employed in categorization experiments) and because it had been employed successfully in a number of previous studies of perceptual learning (Aitken, Bennett, McLaren, & Mackintosh, 1996; McLaren, 1997; McLaren, Leavers, & Mackintosh, 1994; Plaisted & Mackintosh, 1995; Wills & McLaren, 1998; Wills et al., 2004). In the present experiments, stimuli were randomly defined on a trial-by-trial basis, except for an 8×8 configuration of squares. This configuration was entirely predictive of category membership and could appear in any one of the four corners of the stimulus, its position varying randomly from trial to trial. Each stimulus contained just one 8×8 configuration, and for any given participant, only two configurations appeared during training. As an illustration of the relatively cryptic nature of these configurations, the reader is invited to identify the 8×8 configuration that appears in each of the stimuli in the bottom row of Fig. 1.¹ The difficulty of this task for the participants was further increased by the fact that stimuli were presented sequentially, rather than simultaneously.

In addition to examining these *nonobvious* configurations, Experiment 1 also included another (between-participants) condition where the configurations were much more obvious (see Fig. 2, bottom row). Our hypothesis was that unitization would be closer to completion for these obvious configurations. Although the salience reduction process of McLaren–Mackintosh theory predicts a reduction in the rated similarity of obvious configuration-sharing pairs, whether this reduction is observable in Experiment 1 depends on the extent to which it is offset by increases in attention to the obvious configuration, due to its being a perfect predictor of the category label (cf. Rehder & Hoffman, 2005).

Method

Participants and apparatus

Forty-eight students at the University of Exeter took part either for course credits or for £4. They were tested

individually in a quiet cubicle. The training phase and the first similarity judgment task were presented on a PC-compatible computer with a 17-in. CRT monitor running the E-Prime software package (Version 1, Psychology Software Tools, Pittsburgh, PA). The second similarity judgment task (triads) was presented on paper, since this meant that the participants could simply circle the two stimuli that they considered to be most similar to each other.

Stimuli

All the stimuli were viewed from a distance of approximately 0.6 meters and subtended a visual angle of approximately 2.9° .

Training stimuli As is illustrated in Fig. 1, training stimuli were 16×16 black-and-white checkerboards. Each stimulus was generated from a trial-unique arrangement of black and white squares, this arrangement being random with the constraint that exactly half the squares were black. For each training stimulus, a quarter of this randomly defined checkerboard was replaced by one of the fixed 8×8 configurations shown in Fig. 2. Any given participant saw just two configurations during training (nonobvious 1 and 2, nonobvious 3 and 4, obvious 1 and 2, or obvious 3 and 4). These fixed configurations could appear in one of four positions within the stimulus: top right, top left, bottom right, or bottom left. Position of the fixed configuration within the stimulus was manipulated within subjects on a trial-by-trial basis, with each configuration appearing at each location with equal frequency. Each configuration was entirely predictive of the category membership of the stimulus that contained it (e.g., obvious configuration 1 predicted membership in category 1, while obvious configuration 2 predicted membership in category 2).

Test stimuli The test phase employed two types of stimulus pairs: *configuration-sharing pairs* and *noise-sharing pairs*. Configuration-sharing pairs had one of the fixed configurations (Fig. 2) in common, with that configuration appearing in the same location in each member of the pair (but in different, equally frequent locations across trials). The remainder of each stimulus contained a different, random, trial-unique arrangement of black and white squares (i.e., each member of the pair had a different arrangement, and different arrangements were used for each pair). On average, configuration-sharing pairs had 62.5% of squares in common (the 25% of squares making up the configuration, plus, by chance, 50% of the squares making up the noise).

For noise-sharing pairs, 75% of the squares were identical but defined at random for that particular pair. The other 25% of the stimulus for these noise-sharing pairs comprised one of the eight fixed configurations for one stimulus and some more randomly defined noise for the

¹ Reading from left to right, the answers are bottom left, top right, bottom right, top left.

other. Hence, one stimulus in the noise-sharing pair contained a configuration, and the other did not. On average, noise-sharing pairs had 87.5% of squares in common (the 75% of the squares making up the shared noise; in addition, by chance, 50% of the squares contained in the configuration in one stimulus had noise in common with the randomly defined noise in the quadrant of the other stimulus).

Any given participant only saw four configurations during the test phase. If they had been trained on obvious configurations, they were tested of all four obvious configurations (two seen in training, two previously unseen). Similarly, if they had been trained on non-obvious configurations, they were tested on all four nonobvious configurations.

Procedure

Participants were allocated to one of two conditions; participants in the *nonobvious configurations* condition were trained on and subsequently rated stimuli containing nonobvious configurations (Fig. 2, top row), while participants in the *obvious configurations* condition went through the same procedure with stimuli containing obvious configurations (Fig. 2, bottom row). For each participant, two configurations of the appropriate type (i.e., obvious or nonobvious) were used in the training phase, and all four configurations of the appropriate type were used in the test phase. In the nonobvious configurations condition, half the participants saw configurations 1 and 2 in training, while the other half saw configurations 3 and 4 in training (Fig. 2, top row). In the obvious configurations condition, half the participants saw the vertical and horizontal line configurations in training, while the other half saw the concentric squares configurations in training (Fig. 2, bottom row).

At the start of the training phase, participants received instructions informing them that they would see a series of pictures and would be asked to state whether they thought that each one belonged to category A or category B. The instructions reassured participants that, to start with, they would obviously be guessing but that they could learn as they went along. The participants then pressed a key to start the task, and the first training stimulus appeared, displayed centrally against a white background for 2,000 ms,² after which it was replaced by the message “A or B?” (again displayed centrally against a white background). This message remained until the participant made a category decision (pressed either C or M on the PC keyboard, which

had been relabeled “A” and “B,” respectively). Responses made before the “A or B?” prompt appeared were ignored. Once a key had been pressed, appropriate feedback was given on the screen, with information on whether or not the response had been correct and, also, what the correct category had been. Feedback remained on the screen for 1,000 ms, after which the screen was cleared for 500 ms before the next checkerboard appeared and the next trial began. The correct response on each trial was determined by the fixed configuration that appeared in a variable location in the stimulus. The allocation of configurations to category labels was counterbalanced across participants. Stimuli were presented sequentially in blocks of 48 trials, with each configuration appearing 6 times in each of the four locations within the stimulus. Order of stimulus presentation was randomized.

At the end of each training block, participants were told their percent correct in the previous block and were reminded that they should be aiming for 100% correct. If their score in the previous block was less than 90%, they were asked to “keep trying to work out what defines categories A and B.” If they scored 90% or more, they were told that they were doing well but were still reminded that the target was 100%. The participants initiated the next training block by pressing a key. All participants completed 16 blocks of training.

Training was followed by a sequential-pairs similarity judgment task. Participants were told that they were to be shown some pairs of checkerboards one after the other and that they would be asked to rate these pairs on a scale of 1 to 9 according to their similarity, with higher numbers indicating *greater similarity*. They were also told about the sequence of events within each trial. Participants initiated the similarity judgment task by pressing a key. Each trial was signaled by a small black fixation cross appearing for 1,500 ms in the middle of a white screen, at the center of the position at which the checkerboards were to appear. The fixation cross was then replaced by a checkerboard, which remained on the screen for 2,000 ms, followed by a 1,000-ms presentation of a mid-gray square the same size and location as the checkerboard. The mid-gray square was, in turn, replaced by a second checkerboard, which was also presented for 2,000 ms. This second checkerboard was replaced by the message “Rate similarity,” which remained on screen until the participant pressed a number from 1 to 9 (with 9 indicating that the two stimuli were *highly similar*).

The sequential-pairs similarity judgment task comprised 32 randomly ordered trials. Sixteen trials employed configuration-sharing stimulus pairs, and 16 trials employed noise-sharing stimulus pairs. The configuration-sharing trials comprised four presentations of each of the two trained configurations (one presentation in each of the four stimulus locations), plus four presentations of each of the two untrained configurations (one presentation in each

² Unpublished work from our lab indicates that the effects reported in the present article can also be found at stimulus presentation times of 1,000 and 4,000 ms.

of the four stimulus locations). Correspondingly, the noise-sharing trials comprised four presentations of each of the two trained configurations (one presentation in each of the four stimulus locations), plus four presentations of each of the two untrained configurations (one presentation in each of the four stimulus locations). In noise-sharing pairs, the checkerboard containing the configuration was always presented first.

The sequential-pairs similarity judgment task was followed by the simultaneous similarity judgment task. Each trial of this task involved the simultaneous presentation of three checkerboard stimuli at the vertices of an invisible equilateral triangle. Participants were asked to draw a ring around the two checkerboards that looked most similar to each other. They were told that they need not think about their decision too much and should go with their initial judgment. Each stimulus triad was composed of a configuration-sharing stimulus pair and a noise-sharing pair. Specifically, labeling the three stimuli appearing in the triad as stimuli 1, 2, and 3, Stimulus 1 and 2 shared a configuration but were otherwise composed of different noise; stimuli 2 and 3 shared the noise, with stimulus 2 containing a configuration and stimulus 3 having that configuration replaced by more noise. Given that noise-sharing stimulus pairs had more squares in common overall than did configuration-sharing pairs (see the Stimuli section, above), responses for untrained configurations in this triad task were expected to predominately be that the two stimuli of the noise-sharing pair were the most similar. However, unitization of the configuration might be able to partially overcome this difference in overall similarity and increase in the likelihood that the two stimuli of the configuration-sharing pair were judged the most similar. One design aspect of the triad test was that even obvious untrained configurations were expected to predominately result in a noise-sharing response. Hence, if categorization training does increase the similarity of obvious configuration-sharing pairs, the triad task should be well-placed to detect that increase.

The triad similarity judgment task comprised 16 randomly ordered trials, with each of the four configurations seen by any given participant (two trained, two untrained) being presented once in each of the four quadrants of the stimulus.

Results

All tests of significance reported in the present article employed ANOVAs (unless otherwise stated) and were assessed against an α of .05. Huynh–Feldt corrections for nonsphericity were applied where appropriate. For brevity, where tests contain more than one factor, only the F -ratios for significant effects are reported. A complete report of all

multifactor tests can be found in the [Supplementary Materials](#).

Categorization accuracy as a function of training is shown in Fig. 3a. There was a main effect of configuration type (obvious vs. nonobvious configurations), $F(1, 46) = 79.11$, a main effect of training block, $F(15, 690) = 21.36$, and a significant interaction between these two factors, $F(15, 690) = 11.08$. Both the main effect of training block and the training block \times configuration type interaction remained significant as a polynomial order two (i.e., linear) contrast, $F(1, 46) = 46.40$ and $F(1, 46) = 26.73$, respectively.³

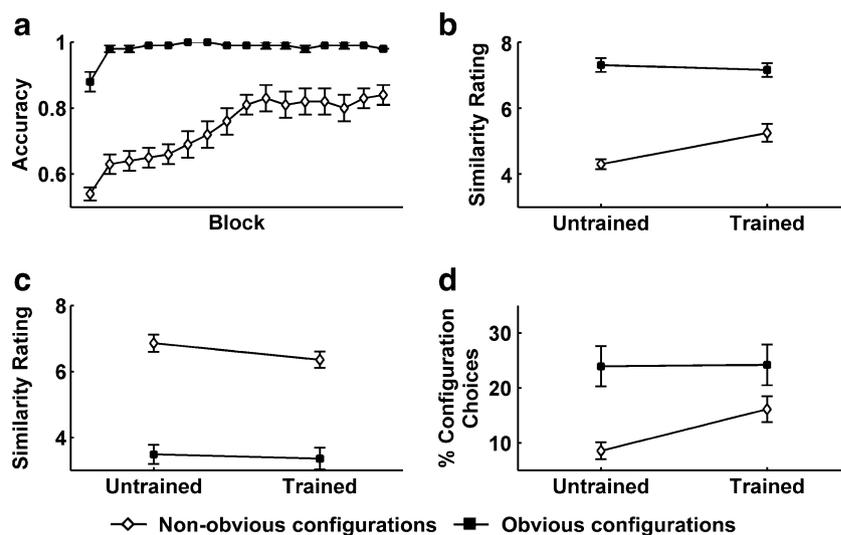
Of principal interest were the effects of categorization training on participants' similarity judgments; these results are shown in Fig. 3b–d. Figure 3b and c show the results for the sequential similarity task; the results for the configuration-sharing pairs (Fig. 3b) and the noise-sharing pairs (Fig. 3c) are shown as separate figures in order to present the observed relationship between training and configuration obviousness clearly.

Figure 3b shows the mean similarity ratings for configuration-sharing pairs in the sequential similarity task as a function of configuration type (obvious vs. nonobvious) and training (trained vs. untrained). There were main effects of training, $F(1, 46) = 6.50$, and of configuration type, $F(1, 46) = 90.28$, and a significant interaction between these two factors, $F(1, 46) = 12.32$. The observed increase in similarity was significant for the nonobvious configurations ($M = 0.95$), $t(23) = 3.26$, while the observed decrease in similarity for obvious configurations was not significant ($M = -0.15$), $t(23) = 1.30$.

Figure 3c shows the corresponding similarity ratings for the noise-sharing pairs in the sequential similarity task. While there was a significant effect of configuration type (obvious vs. nonobvious), $F(1, 46) = 84.32$, neither the main effect of training (trained vs. untrained) nor the interaction was significant. Considering the two configuration types separately, no significant effect of training was revealed in either the obvious configuration condition ($M = -0.14$), $t(23) = 0.77$, or the nonobvious configuration condition ($M = -0.50$), $t(23) = 1.34$. The direction of the main effect of configuration obviousness (i.e., obvious less similar than nonobvious) in the noise-sharing pairs test was to be expected, given that one stimulus of the obvious noise-sharing pair contained a very noticeable configuration, while, in the other stimulus of the pair, that highly noticeable configuration was replaced by noise. This should make the two members of the pair seem rather different. In contrast, the nonobvious configuration was presumably less

³ For further discussion of polynomial contrasts within ANOVAs and a list of the coefficients employed, see Winer, Brown, and Michels (1991).

Fig. 3 Results of Experiment 1. **a** Categorization accuracy during training. **b** Similarity ratings for configuration-sharing pairs in the sequential similarity judgment task. **c** Similarity ratings for noise-sharing pairs in the sequential similarity judgment task. **d** Percentage of “configuration” choices in the triads similarity judgment task. Error bars are one standard error



noticeable, and so the difference between members of the pair was less marked and, hence, the similarity rating was higher.

Figure 3d shows the percentage of occasions on which the members of the configuration-sharing pair were considered to be the most similar in the triad similarity judgment task. There was a main effect of configuration type (obvious vs. nonobvious), $F(1, 46) = 84.32$, a main effect of training (trained vs. untrained), $F(1, 46) = 5.05$, and a significant interaction, $F(1, 46) = 4.40$. The observed increase in percentage was significant for the nonobvious configurations, ($M = 7.55$), $t(23) = 3.03$, while the effect for the obvious configurations was not significant ($M = 0.26$), $t(23) < 1$.

Discussion

In two of our three measures (simultaneous triads and configuration-sharing sequential pairs), category training significantly affected performance. In both cases, this effect was modulated by the type of configuration employed (obvious vs. nonobvious). In the triad task, categorization training increased the likelihood that participants in the nonobvious configurations condition would rate the configuration-sharing stimuli as most similar, rather than the two stimuli that were actually most similar overall (the noise-sharing pair). In the sequential similarity judgment task, mean similarity ratings for configuration-sharing pairs increased as a result of training on nonobvious configurations. Neither effect was significant for the obvious configurations. The fact that the configurations could appear, with equal probability, in any quadrant of the stimulus, means that these changes cannot be explained simply in terms of participants' directing attention toward a fixed spatial location within the stimulus.

The results of Experiment 1 indicate that categorization training led to an increase in the subjective similarity of

configuration-sharing stimuli, but only when the configurations were nonobvious. This pattern of changes in similarity is predicted by the unitization process of the McLaren–Mackintosh theory, but this process also uniquely predicts that increases in the subjective similarity of nonobvious configuration-sharing stimuli should also be observed when the stimuli are simply exposed, rather than actively categorized. This prediction was examined in Experiment 2.

As was discussed in the Introduction, the McLaren–Mackintosh theory also has a salience reduction process, which leads to the prediction that exposure reduces the subjective similarity of *obvious* configuration-sharing stimuli. This effect was not observed in Experiment 1, possibly due to the fact that the configurations were the only aspect of the stimulus that reliably predicted category membership. Under certain accounts (e.g., Kruschke et al., 2005) overt attention would be predicted to become directed to these outcome-predicting configurations, potentially offsetting the effects of the salience reduction process hypothesized by the McLaren–Mackintosh theory. In Experiment 2, the configurations predicted neither category membership nor any other differential outcome. Hence, in Experiment 2, one might expect to observe that exposure leads to a decrease in the rated similarity of obvious configuration-sharing stimulus pairs.

Experiment 2

In most respects, Experiment 2 was the same as Experiment 1, with the test phase being identical. In the training phase, the type of stimuli and amount of exposure to them were also identical to those in Experiment 1. The only way in which Experiment 2 differed from Experiment 1 was that, rather than learn the category assignments of the training stimuli, participants were asked to simply rate each training stimulus for “attractiveness.” Hence, the configurations

present during the training phase no longer predicted any experimenter-defined outcomes.

Method

Participants, apparatus, and stimuli Forty-eight students at the University of Exeter took part either for course credits or for £4. No participant took part in more than one experiment in this study. The apparatus was the same as in Experiment 1, and stimuli were created in the same manner as they were in Experiment 1.

Procedure Prior to starting, participants read a set of instructions informing them that they would be shown some pictures and would have to rate each one on a scale of 1 to 9 according to its attractiveness, with higher numbers meaning *more attractive*. This task was presented within the context of rating the attractiveness of a large set of small abstract paintings (in reality, our checkerboard stimuli).

The participant pressed a key to start the task, and the first stimulus appeared, displayed centrally against a white screen, for 2,000 ms, after which it was replaced by the message “How attractive?” also displayed centrally against a white screen. This message remained until the participant pressed a number from 1 to 9. In this experiment, no feedback was given, so the screen just cleared for 500 ms before the next checkerboard appeared and the next trial began.

After 48 checkerboard patterns had been presented, a message appeared thanking participants for their ratings and telling them to press the space bar to begin the next block. Once all 16 blocks were completed, a message appeared informing participants that the first part of the experiment was over and that they should inform the experimenter. The test phases of the experiment were identical to those in Experiment 1.

Results

Mean attractiveness ratings are shown in Fig. 4a. There was a significant effect of exposure block, $F(15, 690) = 2.04$, and a significant interaction between exposure block and configuration type (obvious vs. nonobvious), $F(15, 690) = 2.18$. Both the effect of block and the block \times configuration type interaction remained significant as a linear contrast, $F(1, 46) = 8.81$ and $F(1, 46) = 6.64$. Configuration type was not significant as a main effect, $F(1, 46) = 1.15$.

As in Experiment 1, it was the participants' similarity judgments that were of principal interest. Figure 4b shows the mean similarity ratings for configuration-sharing pairs in the sequential similarity task as a function of configuration type (obvious vs. nonobvious) and exposure (ex-

posed vs. nonexposed). There was a main effect of configuration type, $F(1, 46) = 18.93$, and a significant interaction between exposure and configuration type, $F(1, 46) = 7.33$. Exposure was not significant as a main effect ($F < 1$). The observed reduction in similarity as a result of exposure was significant for the obvious configurations ($M = -0.59$), $t(23) = 2.59$, while the observed increase in similarity for the nonobvious configurations was not significant ($M = 0.36$), $t(23) = 1.35$.

Figure 4c shows the corresponding similarity ratings for the noise-sharing pairs in the sequential similarity task. While there was a significant effect of configuration type (obvious vs. nonobvious), $F(1, 46) = 39.82$, neither the main effect of exposure (exposed vs. nonexposed) nor the interaction was significant. Considering the two configuration types separately, there was no significant effect of exposure in the obvious configuration condition ($M = -0.22$), $t(23) = 0.95$, or in the nonobvious configuration condition ($M = 0.26$), $t(23) = 0.94$.

Figure 4d shows the mean percentage of trials on which the members of the configuration-sharing pairs were considered to be the most similar in the simultaneous triad similarity judgment task. There was a main effect of configuration type (obvious vs. nonobvious), $F(1, 46) = 26.56$, and a significant interaction between exposure and configuration type, $F(1, 46) = 11.33$. Exposure was not significant as a main effect ($F < 1$). The observed increase in percentage for the nonobvious configurations was significant ($M = 5.21$), $t(23) = 3.05$; the observed decrease in percentage for the obvious configurations was not significant ($M = -2.86$), $t(23) = 1.70$.

Discussion

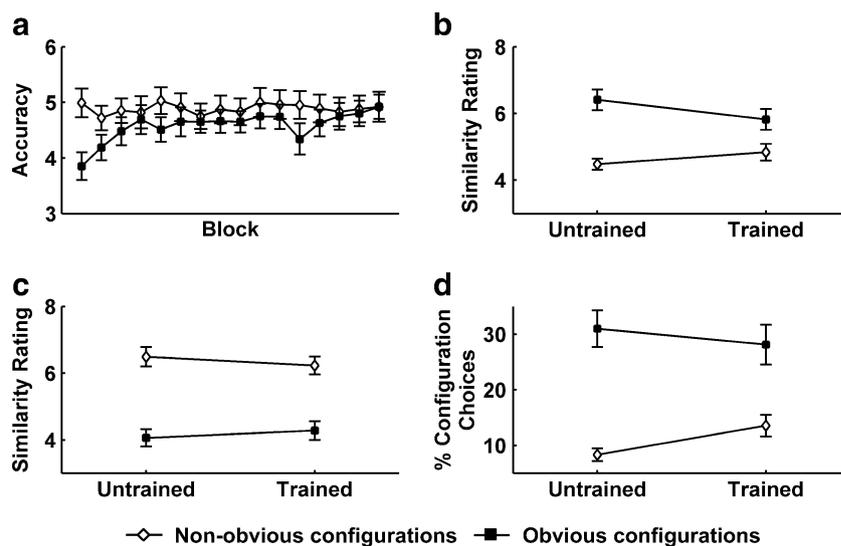
In two of our three measures (simultaneous triads and configuration-sharing sequential pairs), exposure significantly affected performance. In both cases, this effect was modulated by the type of configuration employed (obvious vs. nonobvious), with reductions in similarity for obvious configurations and increases in similarity for nonobvious configurations.

In the triad task, exposure significantly increased the likelihood that participants in the nonobvious configuration condition would rate the configuration-sharing stimuli as most similar, rather than the two stimuli that were most similar overall (the noise-sharing pair). Exposure produced no significant effect in the obvious configuration condition. These results are comparable to those seen in Experiment 1, in which the configurations were familiarized through category training, rather than through exposure, as they were in the present experiment.

In the sequential similarity judgment task, mean similarity ratings for configuration-sharing pairs decreased as a

Fig. 4 Results of Experiment 2.

a Attractiveness ratings during exposure. **b** Similarity ratings for configuration-sharing pairs in the sequential similarity judgment task. **c** Similarity ratings for noise-sharing pairs in the sequential similarity judgment task. **d** Percentage of “configuration” choices in the triads similarity judgment task. Error bars are one standard error



result of exposure to obvious configurations and increased as a result of exposure to the nonobvious configurations. This pattern of numerical trends is broadly comparable to that seen in Experiment 1, but the pattern of significant effects is different. In Experiment 1, the increase in similarity for nonobvious configurations was significant, while the decrease in similarity for obvious configurations was not significant. In Experiment 2, the decrease for obvious configurations was significant, but the increase for nonobvious configurations was not. This raises the question of whether the configuration-sharing sequential similarity results of Experiments 1 and 2 are reliably different.⁴

In order to further investigate this question, we combined the data from Experiments 1 and 2 and performed the same analyses as those performed individually for each experiment, but with familiarization type (categorization vs. attractiveness rating) as an additional between-subjects factor. The hypothesis that the pattern of results is reliably different in the two experiments would gain support from a significant three-way interaction: familiarization type (categorization vs. attractiveness rating) \times familiarization amount (seen vs. unseen configurations) \times configuration type (obvious vs. nonobvious). This interaction does not approach significance, $F(1, 92) < 1$. The corresponding three-way interaction is also not significant for the noise-sharing pairs, $F(1, 92) < 1$, or for the triad task, $F(1, 92) < 1$. In these combined analyses, the interaction between configuration type (obvious vs. nonobvious) and familiarization amount (seen vs. unseen configurations) remains significant for the configuration-sharing pairs, $F(1, 92) = 18.97$, remains significant for the triad task, $F(1, 92) = 13.23$, and remains nonsignificant for the noise-sharing

pairs, $F(1, 92) = 2.38$. These analyses are presented in full in the [Supplementary Materials](#).

Overall, the results of Experiment 2 suggest that exposure increases the similarity of stimuli sharing a nonobvious configuration to a greater extent than it increases the similarity of stimuli sharing an obvious configuration. In the triad task, this is manifested as a significant increase in the nonobvious configuration condition and a nonsignificant decrease in the obvious configuration condition. In the sequential similarity task, it is manifested as a significant decrease in the obvious configuration condition and a nonsignificant decrease in the nonobvious configuration condition. Neither task provides evidence for an increase in similarity for obvious configurations. Taken together, the results of Experiments 1 and 2 are compatible with the unique prediction of the unitization process of the McLaren–Mackintosh theory that unitization of a repeatedly presented set of components should occur whether or not the set of components is diagnostic of some category.

The noise-sharing trials of the sequential similarity task failed to show any effects of category training (Experiment 1) or of exposure (Experiment 2) for either the nonobvious or the obvious configurations. The effects predicted by the McLaren–Mackintosh model are present as numerical trends, but these trends are not statistically reliable. It is possible that the noise-sharing trials would have provided a better and more sensitive measure if the stimuli making up the noise-sharing pair had contained different configurations (rather than one stimulus containing a configuration and the other containing no configuration, as was actually the case).⁵ Resolution of this issue awaits future research.

The salience reduction process of the McLaren–Mackintosh model predicts that the rated similarity of obvious

⁴ A difference in the pattern of significances does not in itself provide evidence for a reliable difference between experiments (see Gelman & Stern, 2006).

⁵ This alternative design was suggested by an anonymous reviewer.

configuration-sharing pairs should be reduced as a result of exposure. The absence of such an effect in Experiment 1 could be explained by the assumption that the salience reduction process is offset by the increase in attention the obvious configurations receive as a result of them perfectly predicting the outcome. This offsetting increase in attention should not have occurred in Experiment 2, because the configurations did not predict any specific outcome. One might, therefore, have expected, under the McLaren–Mackintosh salience reduction process, to see a significant reduction in rated similarity for obvious configuration-sharing pairs as a result of exposure in Experiment 2. We did, in fact, observe this significant decrease once in Experiment 2 (see Fig. 4b), but the evidence is somewhat equivocal, in two regards. First, while the reduction was significant in Experiment 2 (Fig. 4b) and not significant in Experiment 1 (Fig. 3b), the cross-experiment statistical analysis presented above failed to show any significant effect of experiment; such an effect might have been expected to occur. Second, one might also have expected to see a significant reduction for obvious configurations in the simultaneous similarity task of Experiment 2; although the trend was in the correct direction, the observed decrease was not statistically significant.

In Experiment 3, we investigated the salience reduction process further, using an additional methodology. Under the assumption that eye gaze provides an estimate of the salience of stimulus components (with higher salience components attracting more attention and, hence, more of the dwell time), we predicted that exposure to obvious configurations should reduce dwell time to these configurations, because the salience reduction process can be predicted to dominate for obvious configurations (because they unitize more quickly than nonobvious configurations). In contrast, exposure to nonobvious configurations should lead to increased dwell time to these configurations. This is under the assumption that unitization of these nonobvious configurations will lead to an increase in the number of representational elements they activate, which will, in turn, increase their salience and, hence, increase the extent to which they attract attention.

The theoretical position we have taken throughout this article indicates that these predicted effects of unitization and salience reduction might be expected to be overlaid by the effects of outcome predictiveness on visual attention. More specifically, exposure to configurations via categorization training should lead to an increase in the attention the configurations receive, because they perfectly predict the category label. In contrast, exposure to the configurations via attractiveness rating should lead to a reduction in the attention the configurations receive, because they do not predict any specific outcome.

Experiment 3

We used an eye-tracking methodology to investigate the effects of unitization on overt attention. Previous evidence suggests that increased attention to predictive stimulus components is reflected in eye gaze (Kruschke et al., 2005; Wills, Lavric, Croft, & Hodgson, 2007), but the relationship between unitization and eye gaze has so far remained unexamined. In the present experiment, we tracked participants' eye movements throughout training phases that were modified versions of those used in Experiments 1 and 2. The stimuli were enlarged so that the quadrant of the stimulus into which a participant's gaze fell could be more accurately ascertained. Participants received fewer learning trials, leading to a shorter training phase, in order to keep the time for which eye-tracking equipment was to be worn to 35 min (a time that had previously been found to be a sensible upper limit for participants' comfort with our particular apparatus). Each stimulus was displayed for longer than in Experiments 1 and 2 in order to allow more eye movements to be collected on each trial. Following the training phase, during which eye movements were monitored, participants completed a triad task like that in Experiments 1 and 2. Eye movements were not collected during this test phase.

Method

Participants Forty-one students from the University of Exeter took part for course credits or £4. Of these, only 32 produced eye-tracking data that could be effectively analyzed; the remainder of participants had to be discarded due to calibration and data acquisition problems. This is a common problem with eye-tracking studies of categorization (see, e.g., Kruschke et al., 2005).

Stimuli The stimuli were checkerboards, created in the same way as those described for Experiments 1 and 2. However, their size was increased by approximately a factor of two. In the learning phase, in which stimuli were presented on a computer monitor, checkerboards were 55 mm across (subtending approximately 5.25° of visual angle). In the test phase, in which the stimuli were presented on paper, the checkerboards were 46 mm across (approximately 5.85° of visual angle).

Apparatus Eye movements were recorded using an Eye-Link II system (Sensorimotoric Systems/SR Research), a video-based eyetracker with a head movement compensation system. The sampling rate was set at 500 Hz. Participants were seated at a comfortable viewing distance from the display monitor, approximately 60 cm away. A standard PC keyboard was placed in front of participants

for them to indicate their responses. Pupil position was monitored (right eye only) via a miniature infrared CCD video camera mounted on an adjustable headband. The eyetracker was connected to a Dell Celeron PC. A second Dell Celeron PC presented stimuli via a 14-in. monitor and collected the participants' responses via the PC keyboard. Participants sat approximately 60 cm from the monitor. Stimulus presentation and response collection were implemented through the E-Prime software package (Psychology Software Tools Inc.), and a custom-made connector allowed parallel port communications between E-Prime and the eye-tracking apparatus.

Design and procedure We split participants into four groups, reflecting the crossing of two independent variables (configuration type and familiarization type), each of which had two levels (obvious configurations vs. nonobvious configurations and category training vs. attractiveness rating). Of the participants who completed the experiment, there were 8 in each group. Participants took part in two consecutive phases of experimentation: a familiarization phase comprising either a category-learning task or an attractiveness-rating task, followed by a test phase that always comprised a triad similarity-rating task.

The experiment began with an initial calibration of eye movements, and then participants read an instruction screen. Participants were also told to try to concentrate on the stimuli and to be as observant as possible. They then pressed the space bar to start the experiment.

In both conditions, the familiarization phase comprised 240 stimuli, presented in 30 blocks of 8 stimuli. Each trial began with a drift correction, during which participants fixated a centrally presented cross until the recalibration was complete (typically, around 1,000–2,000 ms). The fixation cross was then immediately replaced by the stimulus, which was displayed in the center of the screen for a fixed 3,000 ms. The stimulus was then replaced by a response prompt, also displayed centrally. The prompt was “A or B?” in the categorization condition and “How attractive?” in the attractiveness-rating condition. Participants then made a response using the appropriate keys. In the categorization conditions, this involved pressing either C or M on the PC keyboard, which had been relabeled “A” and “B,” respectively. In the attractiveness-rating condition, the response involved pressing a key from 1 to 9, with higher numbers indicating *more attractive*. There was no time limit for these responses.

In the categorization condition, the participant's response was followed immediately by a centrally displayed feedback message (“CORRECT!” in blue for a correct answer and “INCORRECT” in red for an incorrect answer). This message was displayed for a fixed 1,000 ms. The next trial immediately followed the feedback message. In the

attractiveness-rating condition, there was no feedback, and the next trial immediately followed the participant's response.

At the end of each block, participants received a message on the screen. In the categorization condition, that message informed participants of their accuracy in that block and emphasized that they should be aiming to score 100% correct. In the attractiveness-rating condition, the message requested that participants kept concentrating on the stimuli. At the end of the fifth block, all participants were asked whether they were comfortable and were reminded that if they wished to have a brief break at the end of this or any block (e.g., to close their eyes for a moment), they could.

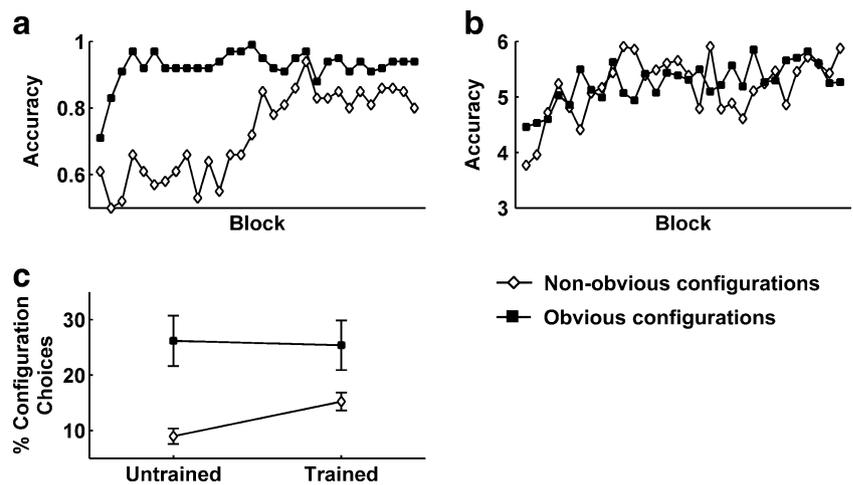
In both conditions, the familiarization phase was followed by a triad test phase, the procedure for which was identical to that employed in Experiments 1 and 2. Eye movements were recorded only during the familiarization phase of the experiment and, within that phase, only during the 3,000-ms period of each trial in which the checkerboard stimulus was presented. All eye track measurements were therefore preresponse.

Results

Categorization accuracy as a function of training is shown in Fig. 5a. There was a main effect of configuration type (obvious vs. nonobvious configurations), $F(1, 14) = 7.12$, a main effect of training block, $F(29, 406) = 4.59$, and a significant interaction between these two factors, $F(29, 406) = 3.2$. Both the main effect of block and the block \times configuration type interaction remained significant as a linear contrast, $F(1, 14) = 15.59$ and $F(1, 14) = 8.61$, respectively. Mean attractiveness ratings as a function of exposure are shown in Fig. 5b. There was a significant effect of exposure block, $F(29, 406) = 3.41$. Neither the main effect of configuration type nor the interaction between configuration type and block was significant.

For the similarity judgment task (Fig. 5c), there was a main effect of configuration obviousness (obvious vs. nonobvious), $F(1, 28) = 8.70$, a main effect of familiarization amount (seen vs. unseen), $F(1, 28) = 5.16$, and an interaction between these two factors, $F(1, 28) = 8.53$. No other effect or interaction approached significance; in particular, there was no significant main effect of, and no significant interaction with, the type of familiarization task (categorization vs. attractiveness rating). The increase in the percentage of “configuration” responses as a result of training on the nonobvious configurations was significant ($M = 6.25$), $t(15) = 4.14$; the trend toward a reduction in configuration responses as a result of training

Fig. 5 Experiment 3, behavioral results. **a** Categorization accuracy during training. **b** Attractiveness ratings during exposure. **c** Percentage of “configuration” choices in the triads similarity judgment task, averaged over familiarization type (categorization vs. attractiveness rating). An average is presented, since there are no significant effects of, or interactions with, familiarization type. Error bars in panel c are one standard error



in the obvious conditions was not significant ($M = -0.13$), $t(15) = 0.42$.

Eye movements were analyzed in terms of configuration dwell proportion (CDP)—the proportion of the static dwell time spent within the checkerboard that was spent within the quadrant containing the 8×8 configuration (i.e., configuration dwell time divided by checkerboard dwell time).⁶ Obvious configurations had a significantly higher CDP than did nonobvious configurations, $F(1, 28) = 4.88$, and participants in the categorization condition had a significantly higher CDP than did participants in the exposure condition, $F(1, 28) = 8.77$. As is illustrated in Fig. 6a, there was a significant interaction between block and configuration obviousness, $F(29, 812) = 4.49$ [this interaction remains significant when analyzed as a linear contrast, $F(1, 28) = 14.71$]. Further analysis revealed a significant increasing linear trend in CDP for nonobvious configurations, $F(1, 15) = 7.96$, and a significant decreasing linear trend in CDP for obvious configurations, $F(1, 15) = 5.02$. As is illustrated in Fig. 6b, there was also a significant interaction between block and familiarization type (categorization vs. exposure), $F(29, 812) = 2.24$, which remained significant when analyzed as a linear contrast, $F(1, 28) =$

6.70. However, the increasing linear trend in the categorization condition was not significant, $F(1, 15) = 2.561$, and the decreasing linear trend in the attractiveness-rating condition was not significant, $F(1, 15) = 2.11$.

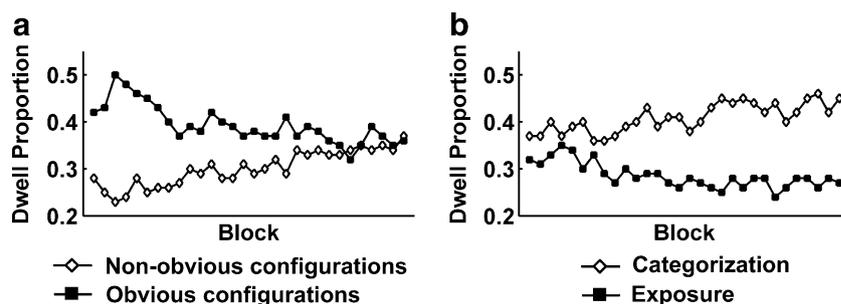
Discussion

The behavioral data in this experiment replicated the findings of Experiments 1 and 2. Analyses of eye gaze indicated that exposure results in an increase in overt attention to nonobvious configurations but a decrease in overt attention to obvious configurations (Fig. 6a). These effects were not significantly modulated by type of exposure (categorization vs. exposure) and, therefore, might be considered to reflect changes in overt attention that are relatively independent of task type. Under the assumption that nonobvious configurations undergo more unitization than do obvious configurations during the experiment, the results shown in Fig. 6a indicate that unitization is associated with an increase in the overt attention to the unitized component.

There was also some evidence of changes in overt attention that were modulated by task type, with categorization and preexposure diverging across training in terms of the proportion of time spent looking at the configuration. However, there was no evidence that this effect of training type was modulated by type of configuration (obvious vs. nonobvious). Hence, the results of the present experiment may indicate the presence of two relatively independent factors determining the change in overt attention to configurations over the course of experience. First, if a configuration is predictive of a category label, categorization training leads to its attracting more overt attention than

⁶ We also analyzed the eye movements using the dependent measure *configuration fixation proportion*. This is the proportion of fixations within the checkerboard that were within the configuration (i.e., the number of configuration fixations divided by the number of checkerboard fixations). The two measures were highly correlated (mean $r = .93$), and the results for configuration dwell time and configuration fixation proportion were extremely similar. In the interests of brevity, only the CDP results are reported.

Fig. 6 Experiment 3, eye movement results. **a** Proportion of dwell time spent in the stimulus quadrant containing the repeated configuration, as a function of training block and configuration type (obvious vs. nonobvious). **b** As in panel a, , but as a function of familiarization type (categorization vs. exposure)



if it serves no explicit purpose (cf. Rehder & Hoffman, 2005). Second, if a configuration is nonobvious, familiarization with that configuration (whether through categorization or through exposure) leads to increases in overt attention to that configuration. If the configuration is obvious, familiarization can lead to a decrease in the overt attention that configuration attracts.

General discussion

In this article, we have focused on one putative mechanism by which people might learn about novel visual configurations. Various associative accounts of learning assume a distributed representation of stimuli and assume that associations may form between those stimulus elements that are repeatedly coactivated (e.g., McClelland & Rumelhart, 1985). One particular model—McLaren and Mackintosh (2000)—was the focus of the present experiments, because it was designed as a model of perceptual learning and because it correctly predicts a number of intriguing perceptual-learning phenomena in humans (e.g., Wills & McLaren, 1998; Wills et al., 2004). Across three experiments, we examined two previously untested predictions of the unitization process of the McLaren–Mackintosh model. The first prediction was that unitization results in the increased similarity of stimuli sharing a reliably recurring component. Experiments 1, 2, and 3 all provide support for this claim. The second prediction was that unitization occurs as a result of exposure and does not depend on the functionality of the exposed components. The results of Experiments 2 and 3 support this prediction. Our data thus suggest that functionality is not a requirement for the development of new feature representations, despite the recent emphasis on functionality in some accounts of representational development (Goldstone, 2000; Schyns et al., 1998; Shiffrin & Lightfoot, 1997).

In the McLaren–Mackintosh model, as in a number of other associative models, unitization results not from functionality, but from the formation of associative links

between co-occurring stimulus elements. The McLaren–Mackintosh model predicts that the formation of these links results in an increase in subjective similarity of stimuli sharing a unitized component, under the assumption that not all elements of a stimulus are sampled on any given presentation but that unitized stimuli can retrieve the nonsampled components via the associative links that have been formed.

Although the present data suggest that functionality is not a requirement for representational development, they also indicate that the extent to which stimulus configurations are functional affects the amount of overt attention they receive. As is illustrated in Fig. 6b, categorization training results in a higher proportion of time spent looking at the part of the stimulus containing the repeated configuration than does preexposure training. Such a result is consistent with the idea that stimulus components that reliably predict an outcome attract more attention than do those that do not (e.g., Kruschke, 2001; Mackintosh, 1975; Wills, 2009). The result illustrated in Fig. 6b is also consistent with experiments that have demonstrated a positive relationship between outcome predictiveness and attention (e.g., Kruschke et al., 2005; Le Pelley, Oakeshott, Wills, & McLaren, 2005; Le Pelley, Reimers, Calvini, Spears, Beesley, & Murphy, 2010; Livesey, Harris, & Harris, 2009; Rehder & Hoffman, 2005; Wills et al., 2007).

Our exploration of eye movements in Experiment 3 further suggests that unitization is associated with an increase in the overt attention attracted by a unitized component. This novel result suggests that unitization may serve an important role in directing overt attention to our environment and that this process favors aspects of the environment whose co-occurrence is currently predicted by the observer, rather than aspects of the environment whose co-occurrence is currently unpredicted.

In addition to the unitization process, the McLaren–Mackintosh theory also posits a salience reduction process, which reduces the activation level of stimulus elements to the extent they are well predicted by other stimulus elements. The elements of a unitized stimulus are well

predicted by the other elements of that stimulus, so there are mechanisms in the model that would tend to reduce the activation level of unitized stimulus elements. This reduction in activation level might, in turn, lead to a reduction in the subjective similarity of stimuli sharing a unitized component (because the elements that they share will have a lowered activity level and, thus, a smaller contribution to similarity as indexed by shared elements). The combination of the unitization process and the salience reduction process means that the McLaren–Mackintosh model predicts that exposure leads to an initial increase in the similarity of nonunitized configuration-sharing pairs (through unitization, which results in an increase in the number of active representational elements), but the model also predicts that, as unitization approaches completion, the similarity of configuration-sharing pairs should be reduced (due to the reduction in the activity level of the unitized elements via the salience reduction process).

In order to investigate the predictions of this salience reduction process, we examined the effect of exposure on stimuli containing configurations that were obvious from the outset (see Fig. 2 for examples). Our hypothesis was that these configurations would unitize quickly and, hence, exposure to these configurations should reveal mainly a salience reduction process (and that, therefore, the similarity of obvious configuration-sharing pairs should be reduced as a result of exposure). No such reduction was observed in Experiment 1, but the absence of an effect in Experiment 1 is explicable in terms of the hypothesized reduction in salience being offset by an increase in overt attention to the configuration, due to it being the only component of the stimulus that predicts the category label. The presence of such an increase in overt attention is consistent with the results of the eye-tracking measures reported in Experiment 3.

In Experiment 2, some evidence was found in support of the salience reduction process, since we observed a significant reduction in the rated similarity of obvious configuration-sharing pairs as a result of exposure. The predictions of the salience reduction process were further supported in Experiment 3, where it was observed that overt attention to obvious configurations, as measured by eye gaze, were significantly reduced during exposure. This effect was opposite to that observed for nonobvious configurations, where overt attention increased.

The process of unitization assumed by the McLaren–Mackintosh model has much in common with the sorts of processes assumed to underlie unsupervised statistical learning; the model assumes people are able to track the co-occurrence of stimulus elements through observation, and studies of statistical learning provide evidence that they are indeed able to do this (e.g., Fiser & Aslin, 2001). One difference is that a recent study of unsupervised statistical

learning indicates that what people learn may be better captured within a Bayesian framework than within a simple associative model (Orban, Fiser, Aslin, & Lengyel, 2008). The extent to which models such as McLaren–Mackintosh are in need in revision in light of such results is an important topic for future research. One might also be tempted to draw parallels between unsupervised statistical learning and situations where people are shown to be able to categorize in the absence of feedback—for example, free classification (e.g., Fried & Holyoak, 1984; Milton, Longmore, & Wills, 2008). On the other hand, the fact that exposure and free classification sometimes have opposite effects (Wills & McLaren, 1998) suggests that the intent to categorize critically affects the processes engaged during unsupervised exposure, with intentional unsupervised categorization perhaps being a form of unsupervised clustering (Love, Medin, & Gureckis, 2004), while unsupervised exposure without the intent to categorization is perhaps better characterized as the learning of stimulus component co-occurrences (Fiser & Aslin, 2001; McLaren & Mackintosh, 2000). The present experiments are consistent with the idea that learning these co-occurrences increases the subjective similarity of stimuli sharing a recurring configuration via the direction of overt attention to those configurations—a process that also seems to occur during, and might facilitate, supervised category learning.

References

- Aitken, M. R. F., Bennett, C. H., McLaren, I. P. L., & Mackintosh, N. J. (1996). Perceptual differentiation during categorization learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 43–50.
- Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, 11, 63–65.
- Donders, F. C. (1969). On the speed of mental processes. In W. G. Koster (Ed.), *Attention and performance II* (pp. 412–431). Amsterdam: North-Holland. Original work published 1868–1869.
- Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, 57, 94–107.
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12, 499–504.
- Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology*, 10, 234–257.
- Gelman, A., & Stern, H. (2006). The difference between “significant” and “not significant” is not itself statistically significant. *The American Statistician*, 60, 328–331.
- Gibson, E. J., & Walk, R. D. (1956). The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *Journal of Comparative and Physiological Psychology*, 49, 239–242.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585–612.
- Goldstone, R. L. (2000). Unitization during category learning. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 86–112.

- Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Interdisciplinary Reviews: Cognitive Science*, *1*, 65–78.
- Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, *45*, 812–863.
- Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 830–845.
- Larkin, M. J. W., Aitken, M. R. F., & Dickinson, A. (1998). Retrospective revaluation of causal judgments under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1331–1352.
- Le Pelley, M. E., Oakeshott, S. M., Wills, A. J., & McLaren, I. P. L. (2005). The outcome-specificity of learned predictiveness effects: Parallels between human causal learning and animal conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*, 226–236.
- Le Pelley, M. E., Reimers, S. J., Calvini, G., Spears, R., Beesley, T., & Murphy, R. A. (2010). Stereotype formation: Biased by association. *Journal of Experimental Psychology: General*, *139*, 138–161.
- Livesey, E. J., Harris, I. M., & Harris, J. A. (2009). Attentional changes during implicit learning: Signal validity protects a target stimulus from the attentional blink. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 408–422.
- Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*, 309–332.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298.
- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–188.
- McLaren, I. P. L. (1997). Categorization and perceptual learning: An analogue of the face inversion effect. *Quarterly Journal of Experimental Psychology*, *50A*, 257–273.
- McLaren, I. P. L., Leavers, H. L., & Mackintosh, N. J. (1994). Recognition, categorization and perceptual learning. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 889–909). Cambridge, MA: MIT Press.
- McLaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, *28*, 211–246.
- McLaren, I. P. L., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, *30*, 177–200.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.
- Melchers, K. G., Lachnit, H., & Shanks, D. R. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *Quarterly Journal of Experimental Psychology*, *57B*, 25–53.
- Milton, F. N., Longmore, C. A., & Wills, A. J. (2008). Processes of overall similarity sorting in free classification. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 676–692.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.
- Orban, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences*, *105*, 2745–2750.
- Plaisted, K. C., & Mackintosh, N. J. (1995). Visual search for cryptic stimuli in pigeons: Implications for the search image and search rate hypotheses. *Animal Behaviour*, *50*, 1219–1232.
- Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*, 1–41.
- Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *The Behavioral and Brain Sciences*, *21*, 1–17.
- Sheu, C. F., & Ratcliff, R. (1995). The application of Fourier deconvolution to reaction-time data: A cautionary note. *Psychological Bulletin*, *118*, 285–299.
- Shiffrin, R. M., & Lightfoot, N. (1997). Perceptual learning of alphanumeric-like characters. In R. L. Goldstone, P. G. Schyns, & D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 36, pp. 45–82). San Diego, CA: Academic Press.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.
- Wasserman, E. A., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgment: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, *51B*, 121–138.
- Weisstein, N., & Harris, C. S. (1974). Visual detection of line segments: Object-superiority effect. *Science*, *186*, 752–755.
- Wills, A. J. (2009). Prediction errors and attention in the presence and absence of feedback. *Current Directions in Psychological Science*, *18*, 95–100.
- Wills, A. J., Lavric, A., Croft, G., & Hodgson, T. L. (2007). Predictive learning, prediction errors and attention: Evidence from event-related potentials and eye-tracking. *Journal of Cognitive Neuroscience*, *19*, 843–854.
- Wills, A. J., & McLaren, I. P. L. (1998). Perceptual learning and free classification. *Quarterly Journal of Experimental Psychology*, *51B*, 235–270.
- Wills, A. J., Suret, M. B., & McLaren, I. P. L. (2004). The role of category structure in determining the effects of stimulus preexposure on categorization accuracy. *Quarterly Journal of Experimental Psychology*, *57B*, 79–88.
- Winer, B. J., Brown, D. R., & Michels, K. E. (1991). *Statistical principles in experimental design*. New York: McGraw-Hill.