Journal of Experimental Psychology: Animal Learning and Cognition

Test Performance in Optional Shift and Configural Acquired Equivalence Are Positively Correlated

Sara Bru García, David N. George, and Jasper Robinson

CITATION

Bru García, S., George, D. N., & Robinson, J. (2024). Test performance in optional shift and configural acquired equivalence are positively correlated.. *Journal of Experimental Psychology: Animal Learning and Cognition*. Advance online publication. https://dx.doi.org/10.1037/xan0000384

Journal of Experimental Psychology:
Animal Learning and Cognition

ISSN: 2329-8456

https://doi.org/10.1037/xan0000384

Test Performance in Optional Shift and Configural Acquired Equivalence Are Positively Correlated

Sara Bru García[1], David N. George[2], and Jasper Robinson[1]
[1] School of Psychology, University of Nottingham
[2] School of Psychology, University of Hull

In two experiments, participants completed two computer-based tasks: a configural acquired equivalence procedure and an optional-shift procedure. Both revealed that test performance was positively correlated, even when controlling for nonspecific variables. This finding supports the suggestion that a common mechanism underlies performance in both tasks. Experiment 2 included eye tracking to the stimuli used in the task. We found that participants who attended to the predictive compound elements in the optional-shift training went on to show stronger attentional-set effects in the subsequent test. The relationship between attention and performance is considered by reference to attentional and nonattentional learning theories.

Keywords: acquired equivalence, attentional set, discrimination learning, configural, optional shift

James (1890) suggested that when two similar stimuli are treated differently—for example, when two wines are labeled "claret" and "burgundy," and are drunk in different situations—they will become easier to differentiate. That is, they will acquire distinctiveness. Miller and Dollard (1941) developed this idea and argued that generalization between two stimuli might increase—they will acquire equivalence—if they are both paired with the same response. These effects are also obtained when specific combinations of stimuli require the same response or predict the same outcome (e.g., Coutureau et al., 2002; Honey & Ward-Robinson, 2001, 2002; Honey & Watt, 1998; Iordanova et al., 2007; Ward-Robinson & Honey, 2000). In one such example of configural acquired equivalence, Robinson and Owens (2013) asked people to learn whether four fictitious people liked or disliked two activities. Each person liked and disliked one of the activities, and each activity was liked and disliked by two people. In one of the counterbalanced conditions, Alice and Charlotte both liked tennis and disliked hockey whereas Beth and Dorothy liked hockey and disliked tennis. Participants needed to learn about the configuration of the person and activity to match them to the correct response. In the second stage of training, one group of participants underwent a reversal of the entire discrimination: Alice and Charlotte now disliked tennis and liked hockey, and Beth and Dorothy now disliked hockey and liked tennis. A second group of participants underwent a partial reversal: The activities that Alice and Dorothy liked and disliked did not change from the first stage, but they did for Beth and Charlotte. The whole reversal group learned the new problem more rapidly than the partial reversal group did. That is, having learned that Alice and Charlotte liked and disliked the same things, participants found it difficult to learn that their likes and dislikes were different. Alice and Charlotte had acquired equivalence as a result of the first stage of training. Thus, the pair of fictitious people were equivalent only in the relationships between their like/dislike and each specific activity.

Duffaud et al. (2007; see also, Kendler et al., 1964; Schwartz et al., 1971) reported an optional-shift procedure that, like configural acquired equivalence, involved learning about combinations of stimuli, with performance transferring from training to test stages. Rats earned food pellets on a two-lever operant discrimination in operant chambers. Audiovisual compounds were used as discriminative stimuli to indicate which of the two levers would be food-reinforced. During Stage-1 training the discrimination had the form: Aw+, Ax+, Bw−, and Bx−, where A and B represent, for example, two auditory elements; w and x represent two visual elements. + indicates that presses to one lever would be food reinforced; − indicates that the lever pressing the other lever would be food reinforced. Thus, in this example, the auditory dimension (A and B) was relevant to the instrumental discrimination and the visual dimension (w and x) was irrelevant. In Stage 2, two new auditory elements (C and D) and two new visual elements (y and z) were introduced. The Stage-2 discrimination had the form: Cy+ and Dz−. In the final, extinction test, rats were presented with a new combination of the Stage-2 elements, Cz and Dy, with responding being biased, respectively, toward the + and − levers. Thus, Stage-1 training

---

Andrew R. Delamater served as action editor.

Sara Bru García https://orcid.org/0000-0002-6406-6864

Jasper Robinson https://orcid.org/0000-0002-1307-5162

Sara Bru García is now at Lynn, Trafalgar House, 5 Fitzalan Place, Cardiff CF24 0ED, United Kingdom.

This work was supported by a British Academy/Leverhulme Small Research Grant (SG163455) to Jasper Robinson. We thank Chan Kwun Fung and Leona Ryan for their help in collecting data and assistance with the eye-tracking equipment.

Correspondence concerning this article should be addressed to Jasper Robinson, School of Psychology, University of Nottingham, University Park Campus, Nottingham NG7 2RD, United Kingdom. Email: jasper.robinson@nottingham.ac.uk

appeared to invoke an attentional set toward the A–D dimension and/or away from the w–z dimension. During Stage 2, C and y were equally good discriminative stimuli for the + lever; and D and z were equally predictive discriminative stimuli for the − lever. However, rats appeared to learn more about the relevant dimension's stimuli, than the irrelevant dimension's stimuli during Stage 2, accounting for the response biases during testing.

Configural acquired equivalence and optional shift have been given different theoretical interpretations with the former conceived of in terms of possible associative/representational changes (e.g., Honey & Ward-Robinson, 2002; Honey et al., 2010; see also, Delamater, 2012) and the latter in terms of changes in attentional set (see, e.g., Duffaud et al., 2007; George et al., 2010; see also, Robbins et al., 1998). These theoretical distinctions seem at odds with the similarities in the general structures of the procedures. Both use multistage training where outcomes are predicted based on specific combinations of stimulus elements. Thus, we sought to investigate the reverse possibility: that configural acquired equivalence and optional shift may be governed by a common mechanism. Without considering any particular mechanism at this stage, we can assume that it will show between-subject variability that will influence performance on both tasks similarly. One consequence of this would be that test performance on the two tasks should be positively correlated; whereas, if different mechanisms governed each task, there would be no such relationship.

The two experiments that we present here supported the common-mechanism interpretation: Both experiments showed positive correlations in performance on configural acquired equivalence (e.g., Honey & Watt, 1998; Robinson & Owens, 2013) and optional-set tasks (e.g., Duffaud et al., 2007; Kendler et al., 1964). People participated in both tasks, whose designs are, respectively, summarized in Tables 1 and 2. Both tasks employed within-subjects designs, allowing performance to be summarized by a single datum and the correlation computed. To aid the comparison of performance, both tasks employed two-stage designs and the stimuli used were counterbalanced across them. Le Pelley et al. (2011; see also Beesley & Le Pelley, 2011) reported evidence of a predictability-driven role for overt attention in performance in tasks, similar to optional shift. Thus, Experiment 2 included eye-tracking measurement to assess the relationship between selective looking during training and the scale of the optional-shift test performance and found similar evidence to Le Pelley et al.

Table 1

Design of the Configural Acquired Equivalence Procedure From Experiments 1 and 2

	Stage 1		Stage 2	
Stimulus	w	x	/	Result: /
A	+	−	+	/
B	−	+	−	/
C	+	−	?	+ predicted
D	−	+	?	− predicted

Note. Some participants predicted which snakes were poisonous or harmless (+/−), during the two-stage configural acquired equivalence procedures, used in Experiments 1 and 2. In both stages, each snake had one of four types of tail (A–D). During Stage 1, only, snakes also had one of two types of skin (w and x), which combined with the four tail types to create eight different snakes (Aw, Ax, Bw, Bx, Cw, Cx, Dw, and Dx). During Stage 2, the snakes had plain skin (i.e., w and x were absent) but they retained their four tail types. One of these snakes would be poisonous (A+) and another would be harmless (B−). Participants were asked to predict whether the two remaining snakes (C and D) were poisonous or harmless, denoted as "?". "/" denotes the absence of a trial type that occurs in other parts of the design. Other participants received the same procedure but with robot stimuli in Experiment 1 and with octopus stimuli in Experiment 2. Participants having the snake task for the configural acquired equivalence had the robot task (Experiment 1), or the octopus task (Experiment 2) for optional shift, and vice versa.

Table 2

Design of the Optional-Shift Procedure From Experiments 1 and 2

	Stage 1		Stage 2		
Stimulus	w	x	y	z	Result: /
A	+	+	/	/	/
B	−	−	/	/	/
C	/	/	+	?	+ predicted
D	/	/	?	−	− predicted

Note. Some of the participants predicted which of the eight robots were dangerous or friendly (+/−). Robots were composed of combinations of four robot images and four robot tones. For some participants, the robot images served as A–D and the robot tones served as w–z; for the remainder, these roles were reversed. In Stage 1, robots' Aw and Ax features were indicative of + and robots' Bw and Bx features were indicative of −. In Stage 2, robots' Cy features were indicative of + and robots' Dz features were indicative of −. Participants were asked to predict whether the two remaining robots (Cz and Dy) were dangerous or friendly, denoted as "?". "/" denotes the absence of a trial type that occurs in other parts of the design. Other participants received the same procedure but with snake stimuli in Experiment 1 and with octopus stimuli in Experiment 2. Participants having robot task for the optional shift had the snake task (Experiment 1), or the octopus task (Experiment 2) for configural acquired equivalence, and vice versa.

Experiment 1

This experiment employed within-subject measures of both configural acquired equivalence (e.g., Honey & Watt, 1998; Robinson & Owens, 2013) and optional shift (e.g., Duffaud et al., 2007; Kendler et al., 1964), whose designs are summarized in Tables 1 and 2, respectively. Participants completed both tasks to allow comparison of performance in the two tasks. In particular, if a common mechanism controls both processes variation in test performance should be positively correlated. The tasks were administered on a computer and involved two different sets of stimuli and scenarios. In one type, participants saw pictures of different snakes and learned to predict which were poisonous (+) and which were harmless (−). In test trials (?), predictions were made but no feedback was given. Snakes were distinguished by their tail types and their skin types. The second task had an identical format but required participants to discriminate robots based on their appearance and their accompanying tones.

Participants varied on whether the snake or robot tasks were used for configural acquired equivalence or for optional shift. Stage-1 training of configural acquired equivalence took the form: Aw+, Ax−, Bw−, Bx+, Cw+, Cx−, Dw−, and Dx+. During Stage-2 training, w and x were now absent and discrimination took the form: A+, B−, C?, and D? That is, A and B were subjects of a nonconfigural discrimination, and generalization of the predictions learned about them to C and D was tested. The expectation was that the prediction for A would transfer relatively well to C; and that the prediction for B would transfer relatively well to D. Stage 1 of the optional-shift task had the form Aw+, Ax+, Bw−, and Bx−. That is, A and B perfectly predicted their outcomes, and w and x were not predictive. Stage-2 discrimination had the form Cy+, Dz−, Cz?, and Dy?. Here C, y, D, and z elements are equally predictive of their outcomes but learning was expected to be biased toward learning about the C and D elements over learning about the y and z elements. This was tested with the Stage-2 Cz? and Dy? trials, where the optional-shift effect (e.g., Duffaud et al., 2007; Kendler et al., 1964) would be seen if Cz and Dy were more predictive, respectively, of + and −.

Method

Participants

Thirty-two students from the University of Nottingham participated (10 male and 22 female, $M_{\text{age}} = 26.06$, $SD = 4.08$, range = 21–34). This sample size was based on the requirement of the experimental design to have equal numbers of participants in each of the subgroups created by stimulus counterbalancing. Participants were informed about the nature and requirements of the task prior to the start of the experimental session and debriefed after it. None of the participants had participated in our other, related experiments and so may be taken to be naïve with respect to the stimuli that we used. The School of Psychology, University of Nottingham's Research Ethics Committee approved the experiment.

Apparatus and Stimuli

The experiment was conducted in a small quiet room in the School of Psychology at the University of Nottingham. Participants were tested individually, sitting in front of a desk, at approximately 50 cm from a computer (iMac, Apple Computers) whose display was 52 (width) × 38 (height) cm. A standard 105-key ISO QWERTY keyboard, with a number pad to the right of the alphabetic keys, was attached to the computer and was placed immediately in front of the computer display. The computer was used to present experimental events and to record keyboard responses during the experiment and employed the Python user interface, PsychoPy (Peirce et al., 2019; Version 1.82.02) (Computer software).

Two sets of cartoon images of snakes and robots were used as stimuli and could be presented on the computer display. The images were produced by Joint Photographic Experts Group (.jpg) files and were irregularly shaped but occupied a 10-cm wide and 8-cm heigh rectangular space. Snakes could have one of four types of tail (pointed, forked, axe shaped, and with a rattle) and one of four types of skin patterns (oval, triangles, spots, and stripes). These features could vary independently and, therefore, there were 16 possible permutations of tail and skin pattern. Four images of robots were used. All stood upright, facing forward but they differed in features such as head and body shape, and coloring. Robots could be accompanied by one of four tones, which differed in features such as pitch and pulsing. Thus, there were 16 possible permutations of robot images and robot tones. The tones were produced by Waveform Audio File Format (.wav) files, created using Audacity (Version 2.3.0) [Computer software], and were presented through a pair of headphones (Panasonic RP-HT225), which participants wore. The white text could also be presented on a gray background. White and gray colors were achieved by setting color channels, respectively, to their full (i.e., 255) and mid-range (i.e., 128) eight-bit values.

Procedure

All participants completed both configural acquired equivalence and optional-shift tasks. Half of the participants received snake stimuli for the configural acquired equivalence task and robot stimuli for the optional-shift task; the remainder received the alternative arrangement. Half of the participants received configural acquired equivalence before optional shift; the remainder received the tasks in the alternative sequence.

Participants read an instruction sheet before beginning. The experimenter left the room after ensuring participants had understood the tasks and returned only to set up the second task, before leaving again until the end of the experiment. During the snake version of the tasks, participants were presented with on-display instructions:

> Imagine yourself in the role of a rainforest tour guide. It is your job to make sure tourists are safe during the duration of the tour. You are about to enter an area densely populated by snakes, some of which are known to be dangerous to humans. It is your task to look at the snakes and learn which ones are poisonous.

That is, participants would learn to anticipate which snakes, differing in their tail and skin pattern, were poisonous and which were harmless. In the robot version of the tasks, participants were given on-display instructions:

> It is the year 2250 and robots have risen against humanity! Fortunately, not all robots present a risk to humans. You will be presented with some robots and robot noises simultaneously. It is your task to learn which robots are dangerous.

That is, participants would learn to anticipate which robots, defined by their appearance and accompanying tone, were dangerous and which were not. The left and right arrow keys and the q and z keys were used, respectively, in the configural acquired equivalence task and optional-shift tasks. Left would indicate that a snake was poisonous or that a robot was evil; right would indicate that a snake was harmless or that a robot was friendly.

Every trial began with the presentation of the text "Get ready!" for 0.5 s. A snake or robot image was then presented in the center of the display with text about the keyboard response requirement for that trial. The robot image would be accompanied by a tone, where this was part of the discrimination. The stimulus and text extinguished on the participant's response, or after 5 s, if none was made. Centrally located feedback text was displayed for 1 s after this, either "Correct!" or "Ooops! That was wrong." If a participant failed to respond, they received the same feedback as on an incorrect response. Feedback was followed by the stimulus and the text stating either that the snake was poisonous/harmless or that the robot was dangerous/friendly for 2 s.

For configural acquired equivalence, Stage 1 included eight individually presented trial types: Aw+, Ax−, Bw−, Bx+, Cw+, Cx−, Dw−, and Dx+ (see Table 1). There were 12, eight-trial blocks, sequenced randomly with the constraint that each trial type occurred once in each block. In the snake version of the task, the four tail types served as A–D, and the striped skin and oval skin served as w and x. The other two types of skin patterns were not used. Eight subgroups were created by counterbalancing the roles of specific stimuli. In the robot version of the task, the four robot tones served as A–D and two of the robot images served as w and x. The two other robot images were not used. There was no intertrial interval. During Stage 2, there were three blocks of four trial types: A+, B−, C?, and D? That is, A+ and B− trials comprised an explicit discrimination and the transfer of this learning was assessed to C? and D?—the means of assessing acquired equivalence. Snake images had one of the four tail types (A–D) during Stage 2, but no skin pattern (w–x). Robot tones (A–D) were presented during Stage 2 but with no robot image (w–x). Participants made predictions about the outcomes of the A–D trials but the feedback was given only on A+ and B− trials. Participants were required to guess the outcome of the C? and D? trials. On these trials, the text "The snake escaped before you could catch it! Your feedback couldn't be delivered this time" or "The robot disrupted the signal! Your feedback couldn't be delivered" was presented in place of the usual feedback. Participants received no indication that they had advanced to Stage 2 trials. Unspecified details of Stage 2 were identical to those of Stage 1.

All four robot images and robot tones, and all four snake skin and snake tail types, were used in the optional-shift tasks. In the robot task, the images served as A–D and the tones as w–z for half of the participants; for the remainder, these roles were reversed. In the snake task, the skin types served as A–D and the tail types as w–z for half of the participants; for the remainder, these roles were reversed. This and other variations of the roles of specific stimuli created eight counterbalanced subgroups. Stage 1 comprised the presentation of four trial types Aw+, Ax+, Bw−, and Bx−. Stage 2 introduced four compound stimuli composed of new elements: Cy+, Dz−, Cz?, and Dy? The four trial types in both stages were repeated once in each of the 12 blocks (i.e., there were 48 trials, in total, in both stages). All unspecified details are identical to those for the configural acquired equivalence task.

Data Treatment and Analysis

The proportion of correct trials per block was computed for each stage of the acquired equivalence and optional-shift tasks. Although trials were intermixed during the tasks, transfer trials in Stage 2 were analyzed separately from test trials that had no feedback. For a test trial to be correct in the acquired equivalence task, it meant participants had transferred their responses to C and D based on the transfer trials provided to the stimulus that had been trained as equivalent during Stage 1 of the task (i.e., C and D are indicative of, respectively, + and −). For a test trial to be correct in the optional-shift task, it meant participants had demonstrated a bias for the dimension established as relevant (A–D) during Stage 1 (i.e., Cy and Dz are indicative of, respectively, + and −). Test trials from both tasks were averaged to obtain a single datum per participant and correlated to determine the relationship between performance in both tasks. Data were collapsed over the counterbalanced subgroups.

Data were analyzed using analysis of variance (ANOVA), one-sample *t* tests, and Pearson's correlations. *t* tests and ANOVAs were tests of two-tailed hypotheses, with an α of .050. One-sample *t* tests were used to evaluate deviation from chance ($\mu = .5$), with data averaged over all trials of an entire experimental stage. Correlations were tests of one-tailed, positive, relationships. These analyses were performed with Jamovi (Version 2.3.28.0) (Computer software). Standardized 90% confidence intervals (CIs) for η_p^2 were computed using the methods described by Kelley (2007) and used his MBESS package (Version 4.9.3) (Computer software) for R. The R scripts were run in RStudio (Version 2023.06.1+524) (Computer software).

Transparency and Openness

Data, analyses, sample stimuli, and scripts for computing η_p^2 CIs can be accessed from https://osf.io/7dvf8/. This experiment was not preregistered.

Results and Discussion

Configural Acquired Equivalence

Data for Stage 1 are summarized in the leftmost panel of Figure 1. Participants acquired the discrimination, $F(11, 341) = 18.5$, $MSE = 0.54$, $p < .001$, $\eta_p^2 = .373$, 90% CI [0.290, 0.414], and performed reliably above chance, $t(31) = 7.72$, $p < .001$, $d = 1.37$, 95% CI [0.87, 1.84]. Discrimination for the transfer trials (A+ and B−) is summarized in the rightmost panel of Figure 1. It increased over the course of training, $F(11, 341) = 8.0$, $MSE = 0.42$, $p < .001$, $\eta_p^2 = .205$, 90% CI [0.122, 0.242] and was reliably above chance, $t(31) = 14.2$, $p < .001$, $d = 2.51$, 95% CI [1.79, 3.21].

The data of the test from Stage 2 are of central interest, here, and are summarized in the rightmost panel of Figure 1. Although the discrimination did not improve over testing, $F(11, 341) = 1.3$, $MSE = 0.14$, $p > .221$, $\eta_p^2 = .040$, 90% CI [0.000, 0.047], it was reliably above chance, $t(31) = 3.85$, $p < .001$, $d = 0.68$, 95% CI [0.29, 1.06]. Thus, these results confirm the reliability of the configural acquired equivalence effect reported in human participants (e.g., Delamater & Joseph, 2000; Robinson & Owens, 2013) and transfer procedure employed by Honey and Watt (1998; see also Ward-Robinson & Honey, 2000).

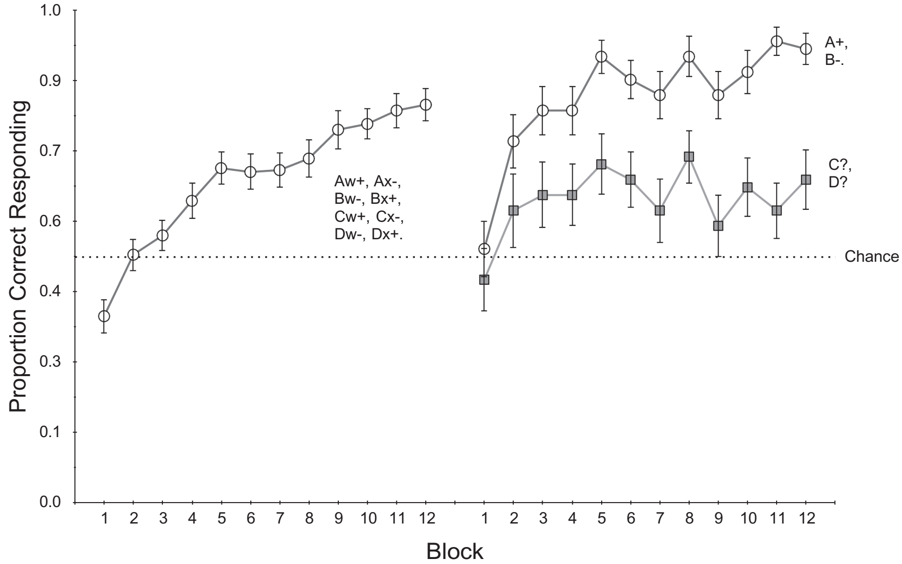
Optional Shift

Data for Stage 1 are summarized in the leftmost panel of Figure 2. Discrimination improved over training, $F(11, 341) = 31.1$, $MSE = 0.46$, $p < .001$, $\eta_p^2 = .501$, 90% CI [0.426, 0.537], and was reliably above chance, $t(31) = 40.2$, $p < .001$, $d = 7.11$, 95% CI [5.31, 8.85]. Stage 2's Cy+ and Dz− training is summarized in the rightmost panel of Figure 2. Again, this improved over the course of training, $F(11, 341) = 9.9$, $MSE = 0.41$, $p < .001$, $\eta_p^2 = .245$, 90% CI [0.158, 0.282] and was reliably above chance, $t(31) = 20.3, p < .001$, $d = 3.60$, 95% CI [2.64, 4.55].

The Cz? and Dy? data from the test of Stage 2 are summarized in the rightmost panel of Figure 2. As with the acquired equivalence test data, there was little apparent improvement in discrimination, $F(11, 341) = 0.6$, $MSE = 0.05$, $p > .828$, $\eta_p^2 = .019$, 90% CI [0.000, 0.012], but of most importance, performance transferred reliably, $t(31) = 3.82$, $p < .001$, $d = 0.675$, 95% CI [0.29, 1.05]. According to many theories of associative learning (e.g., Rescorla & Wagner, 1972) the elements C and y should be equally good, partial predictors of +; and D and z of − (e.g., each element would be $\lambda/2$ for its outcome). Accordingly, there should be no net difference in the expectations of + and − when these elements are recombined in the compounds Cz? and Dy? That is, C's $\lambda/2$ for + would be matched by y's $\lambda/2$ for +, and D's $\lambda/2$ for − would be matched by z's $\lambda/2$ for −. Each compound would partially activate the two outcomes, to an equivalent extent. However, the transfer of + responding from C was greater than the transfer of + responding from y. And/or the transfer of − responding from D was greater than the transfer of responding from z. That is, the Stage-1 discrimination had biased what was learned during Cy + and Dz−, toward the relevant A–D dimension, and/or away from the irrelevant w–z

Figure 1

Means and Their Standard Errors From the Configural Acquired Equivalence Task From Experiment 1



Note. Performance is expressed as the proportion of correct responses, with 0.5 representing chance which is indicated by a horizontal, dashed line. A–D, w–z, +/− and ? represent the format of the trials in the experiment. A–D and w–z represent experimental stimuli. + and − represent outcomes that participants predicted from A–D and w–z. ? indicates that no outcome was presented. The leftmost panel summarizes data from Stage 1, the initial configural discriminations (Aw+, Ax−, Bw−, Bx+, Cw+, Cx−, Dw−, and Dx+). The rightmost panel summarizes data from the transfer trials (A+, B−, circles) and test trials (C?, D?, squares) in Stage 2. Stage 1 and Stage 2 data are blocked over, respectively, eight and three trials.

dimension, demonstrating the optional-shift effect (e.g., Duffaud et al., 2007; Kendler et al., 1964).

Relationship Between Configural Acquired Equivalence and Attentional Set

We looked at the relationship between participants' overall test performance in both tasks of Experiment 1. These data are presented Figure 3 and indicate a positive relationship between test performance on configural acquired equivalence and optional shift, $r(30) = .39$, $p < .013$. This relationship can be readily understood if we assume that a common mechanism governs performance on both tasks. Such a mechanism will show between-subject variability that will transfer to both procedures causing performance to be related.

It is important to note that, other variables, such as motivation or arousal also seem likely to contribute to this correlation and, in principle, could fully account for it. This would mean that our results could not be safely taken as evidence of any relationship in performance between the two types of tests. One solution to this problem is to reexamine the test correlation using a control variable in a partial correlation (see, e.g., Kim, 2015; Yule, 1919, p. 238; van Aert & Goos, 2023). We reasoned that the performance averaged over all Stage-1 trials of both tasks, would serve as the control variable because the influence of nonspecific variables should be the same, here, as in the two tests. That is, not to say that there should not be any correlation between Stage-1 performance and subsequent test performance (cf. Coutureau et al., 2002); rather that the partial correlation can examine any residual correlation between the two tests, free from the influence of Stage-1 performance. The partial correlation uses the three correlations between configural acquired equivalence test performance, optional-shift test performance, and Stage-1 performance; and it subtracts the two correlations between Stage-1 performance and performance in the two tests. This method also requires a modification of the original correlation's degrees of freedom (e.g., Kim, 2015; Weatherburn, 1961). Following these procedures, the correlation, $r(27) = .38$, $p < .017$, retained reliability and its effect size was only marginally smaller than the original correlation's.

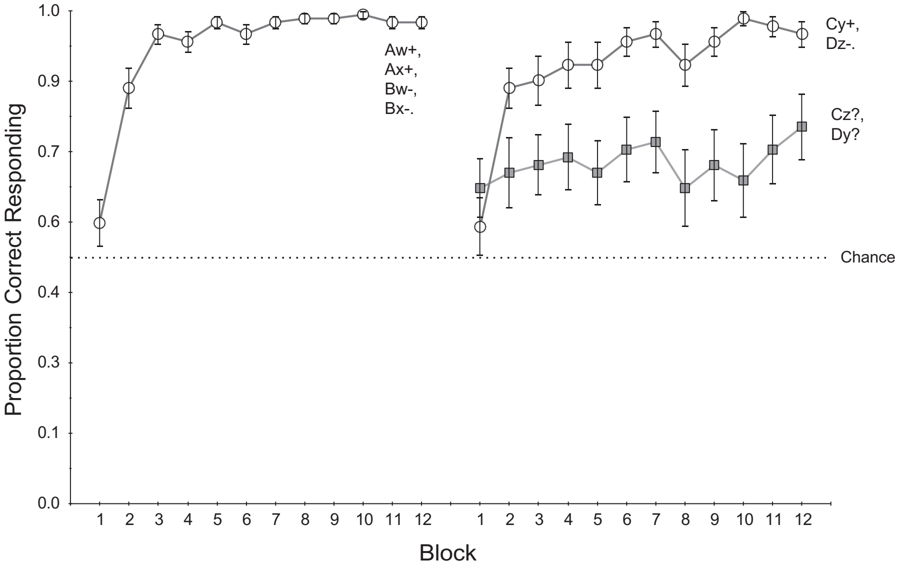
Experiment 2

Experiment 1 demonstrated configural acquired equivalence using stimulus transfer trials (cf., Honey & Watt, 1998; Iordanova et al., 2007; Ward-Robinson & Honey, 2000) and optional shift (e.g., Duffaud et al., 2007; Kendler et al., 1964) in human participants. The results demonstrated a positive relationship between performance in these two tasks that could not be attributed to nonspecific variables. This feature of our findings implies that a mechanism common to both forms of learning is operational, which, in turn, causes performance on them to be related.

Le Pelley et al. (2011; see also Le Pelley & McLaren, 2003) used a learned predictiveness task, similar to our optional-shift task, that uncovered an apparent role for overt attention. Participants received a three-stage procedure in which outcomes could be predicted based

Figure 2

Means and Their Standard Errors From the Optional-Shift Task From Experiment 1



Note. Performance is expressed as the proportion of correct responses, with 0.5 representing chance which is represented by a horizontal, dashed line. A–D, w–z, +/− and ? represent the format of the trials in the experiment. A–D and w–z represent experimental stimuli. + and − represent outcomes that participants predicted from A–D and w–z. ? indicates that no outcome was presented. The leftmost panel summarizes data from Stage 1, the initial Aw+, Ax+, Bw−, and Bx− discriminations. The rightmost panel summarizes data from the transfer trials (A+, B−, circles) and test trials (C?, D?, squares) in Stage 2. Stage 1 and Stage 2 data are blocked over, respectively, eight and three trials.

on the presentation of visual cues that were assembled from separable elements. As in our experiments, predictions were recorded by keyboard responses. During Stage-1 training participants were given trials having the form: Aw+, Ax+, Bw*, Bx*, Cy*, Cz*, Dy+, and Dz+. Notice that, as in the optional-shift design, A–D perfectly predict either outcome + or outcome *, whereas w–z are equally often predictive of + and *. Stage-2 training replaced these outcomes with two new outcomes and had the form: Az$, By£, Cx$, and Dw£. Here elements A–D and w–z are equally predictive of the two new outcomes. During testing (Stage 3) the eight elements were presented in four new compounds, AC, BD, wy, and xz, and participants predicted the $ and £ outcomes. Notice that each compound is composed of two elements that have predicted the same Stage-2 outcome and that, because all elements were equally predictive of their outcome, according to many models of associative learning (e.g., Pearce, 1994; Rescorla & Wagner, 1972), each compound should be judged good predictors of either $ (i.e., AC and xz) or £ (i.e., BD and wy). However, the AC and BD compounds were judged to be better predictors of their outcomes than were wy and xz. The only difference in the treatments of these compounds' elements was the reliability with which they predicted their outcomes in Stage 1. Thus, we might describe Le Pelley et al.'s procedure as a variant of optional shift with new outcomes, rather than new stimulus compounds in the transfer stage (cf. Table 2).

Le Pelley et al. (2011) used an eye tracker to record times spent looking at each of the two elements of their compound stimuli, implying overt attention. In both training stages, their participants biased looking toward the element of each compound that perfectly predicted its outcome during Stage 1, relative to its imperfectly predictive partner. They also found a reliable, positive correlation between this bias to look more at predictive compound elements, during Stage 2, and performance in the subsequent test. They suggested that these results were consistent with a three-part causal process in which highly predictive stimuli promote overt attention, which, in turn, promotes effective learning. The purpose of Experiment 2 was to both replicate the novel findings of Experiment 1 and to include Le Pelley et al.'s measurement of eye tracking. In particular, we wished to examine any positive relationship between overt, selective looking to the predictive training elements and to the test performance.

Method

Participants, Apparatus, Stimuli, and Procedure

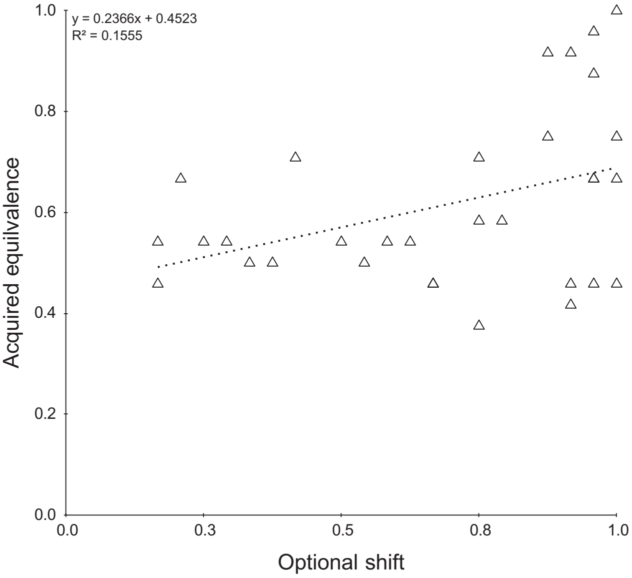
Thirty-two students from the University of Nottingham participated (14 male and 18 female, $M_{\text{age}} = 21.56$, $SD = 2.10$, range = 18–25).

The experiment was run in a larger laboratory than was used in Experiment 1. Participants were located behind a 2-m tall hessian pinboard, intended to promote focus on the task. The experimenter sat quietly behind the pinboard during the experiment. The experiment was run on a Tobii TX300 eye tracker (Tobii Technology, Danderyd, Sweden) with a 51 (width) × 28 (height) cm display, and a display-mounted eye tracker recording recorded gaze at a resolution frequency of 60 Hz. It did not require a chin rest. Participants sat approximately 50 cm from the display.

The snake versions of the two tasks from Experiment 1 were also used in Experiment 2. However, the robot versions of the tasks used in Experiment 1 were unsuitable for Experiment 2 because it

Figure 3

Scattergram of the Relationships Between Performance on the Configural Acquired Equivalence (Ordinate) and Optional-Shift Task (Abscissa) in Experiment 1



Note. Performance is expressed as the proportion of correct responses, with 0.5 representing chance. Lines represent the linear regression of optional-shift performance on the configural acquired equivalence task, whose values are presented in the top left of the figure with R^2, the proportion of variance in performance in acquired equivalence predicted by optional-shift performance.

employed auditory stimuli as A–D, which would not be detected by the eye tracker. These tasks were substituted for an alternative, purely visual task in which cartoon images of octopuses could have one of four different types of eyes and one of four different types of tentacles, which could be combined as they were in Experiment 1 (see Tables 1 and 2). The octopus images were irregularly shaped but were 10 cm wide and 8 cm high. Participants predicted whether each type of octopus would bite or would sting. Before the start of the octopus tasks, participants read a set of instructions asking them to,

> Imagine yourself in the role of a marine tour guide. It is your job to keep tourists safe from all dangerous animals. Your boat is about to enter an area densely populated by octopuses that are known to be dangerous to humans.

The instructions indicated that it was the participants' task to "look at the octopuses and learn which ones can bite you." All unspecified other details were identical to those of Experiment 1.

Two regions of interest (ROIs) were used for each compound stimulus. ROIs had different sizes for the snake and octopus cartoons to accommodate for the differences in the locations of their distinctive features. The ROIs were 3.0 cm × 3.0 cm and 3.5 cm × 3.5 cm, respectively, for the octopuses' eyes and for snakes' tails. The ROIs were, respectively, 9 cm wide and 6 cm high for the octopuses' tentacles and 3.5 cm wide and 3.5 cm high for the snakes' skin. The eye tracker recorded only during image presentation.

Results and Discussion

Configural Acquired Equivalence

Experiment 2's Stage-1 performance is summarized in the leftmost panel of Figure 4. It shows a gradual improvement in performance, $F(11, 341) = 12.7$, $MSE = 0.27$, $p < .001$, $\eta_p^2 = .291$, 90% CI [0.205, 0.331], and above-chance performance, $t(31) = 8.35$, $p < .001$, $d = 1.47$, 90% CI [0.97, 1.98]. The A+ and B− performance of Stage 2 are summarized in the rightmost panel of Figure 4 and show a similar improvement in performance, $F(11, 341) = 15.8$, $MSE = 0.59$, $p < .001$, $\eta_p^2 = .337$, 90% CI [0.252, 0.378]. Again, A+ and B− performance exceeded chance, $t(31) = 26.35$, $p < .001$, $d = 4.66$, 90% CI [3.45, 5.86].

Accuracy data from test trials with C? and D? are summarized in the rightmost panel of Figure 4. Like the test results of Experiment 1, there was a reliable change in performance across testing, $F(11, 341) = 6.8$, $MSE = 0.43$, $p < .001$, $\eta_p^2 = .180$, 90% CI [0.099, 0.215], and of crucial importance, overall performance was reliably above chance, $t(31) = 2.88$, $p < .007$, $d = 0.51$, 95% CI [0.14, 0.87], replicating the results of Experiment 1 (see also, e.g., Honey & Watt, 1998; Iordanova et al., 2007; Ward-Robinson & Honey, 2000).

Because no individual stimulus uniquely predicted the outcome of any given trial, we anticipated no differences in average fixation time to any particular dimension during the acquired equivalence task. A repeated measures ANOVA, with factors of dimension and stage, confirmed no differences in dwell times across stages and yielded no significant main effects or interactions (smallest $p > .098$ for the main effect of stage).

Optional Shift (Task Performance)

Stage-1 optional-shift performance is summarized on the leftmost panel of Figure 5. As in Experiment 1, performance improved promptly, $F(11, 341) = 33.8$, $MSE = 0.41$, $p < .001$, $\eta_p^2 = .521$, 90% CI [0.449, 0.557] and was reliably better than chance, $t(31) = 58.51$, $p < .001$, $d = 10.34$, 95% CI [7.66, 12.86]. Figure 5 (rightmost panel) shows the discrimination to performance with the new compounds Cy+ and Dz− during Stage 2. Performance improved over Stage 2, $F(11, 341) = 13.4$, $MSE = 0.27$, $p < .001$, $\eta_p^2 = .302$, 90% CI [0.216, 0.343] and was above chance, $t(31) = 25.7$, $p < .001$, $d = 4.55$, 90% CI [3.37, 5.73].

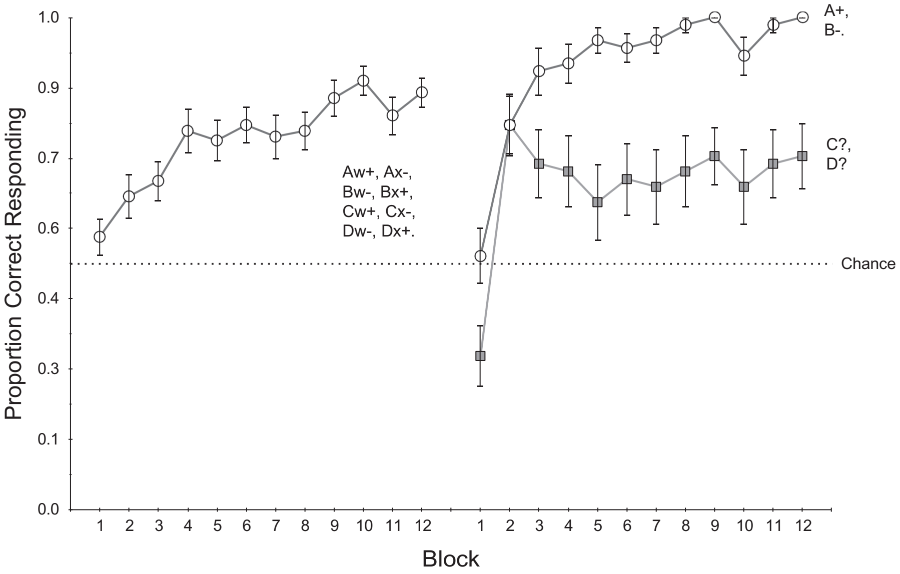
Accuracy to test trials Cz? and Dy? is summarized in the rightmost panel of Figure 5. As in Experiment 1, this confirmed participants' bias toward the A–D dimension that had been predictive during Stage-1 training and/or away from the w–z dimension, replicating the optional-shift effect (cf., e.g., Duffaud et al., 2007; Kendler et al., 1964), $t(31) = 5.88$, $p < .001$, $d = 1.04$, 90% CI [0.60, 1.47]. There was no reliable change in performance, $F(11, 341) = 0.5$, $MSE = 0.03$, $p > .383$, $\eta_p^2 = .019$, 90% CI [0.000, 0.012].

Relationship Between Configural Acquired Equivalence and Attentional Set

As in Experiment 1, participants' performance during test trials in the configural acquired equivalence and optional-shift tasks was correlated, $r(30) = .54$, $p < .001$. These data are presented in Figure 6. Thus these findings further support the notion that attentional set and configural acquired equivalence can be accommodated by the same process, with individual variation in one task covarying with

Figure 4

Means and Their Standard Errors From the Configural Acquired Equivalence Task From Experiment 2



Note. Performance is expressed as the proportion of correct responses, with 0.5 representing chance which is indicated by a horizontal, dashed line. A–D, w–z, +/− and ? represent the format of the trials in the experiment. A–D and w–z represent experimental stimuli. + and − represent outcomes that participants predicted from A–D and w–z. ? indicates that no outcome was presented. The leftmost panel summarizes data from Stage 1, the initial configural discriminations (Aw+, Ax−, Bw−, Bx+, Cw+, Cx−, Dw−, and Dx+). The rightmost panel summarizes data from the transfer trials (A+, B−, circles) and test trials (C?, D?, squares) in Stage 2. Stage 1 and Stage 2 data are blocked over, respectively, eight and three trials.

individual variation in the other. Again, this correlation was retained when performance from all Stage-1 trials of both tasks was averaged and used as a control for nonspecific variables in a partial correlation, $r(27) = .44$, $p < .007$.

Optional Shift (Eye Tracking)

Figure 7 indicates that participants looked, selectively toward the relevant dimension (A–D), relative to the irrelevant dimension (w–z), in all stages. An ANOVA with factors of dimension and discrimination revealed a significant main effect of dimension, $F(1, 31) = 22.4$, $MSE = 2.17$, $p < .001$, $\eta_p^2 = .419$, 90% CI [0.190, 0.570]. The means were not parallel over the three discriminations but neither the Relevant/Irrelevant × Discrimination interaction, nor the main effect of discrimination was reliable, smallest $p > .099$. Dwell times to the relevant dimension exceeded those to the irrelevant dimension on all three of the parts of the discrimination, smallest, $t(31) = 3.30$, $p < .003$, $d = 0.58$, 90% CI [0.20, 0.95]. These findings mirror those reported by Le Pelley et al. (2011) for a similar learned predictiveness procedure (see also, Beesley & Le Pelley, 2011; Griffiths & Le Pelley, 2019).

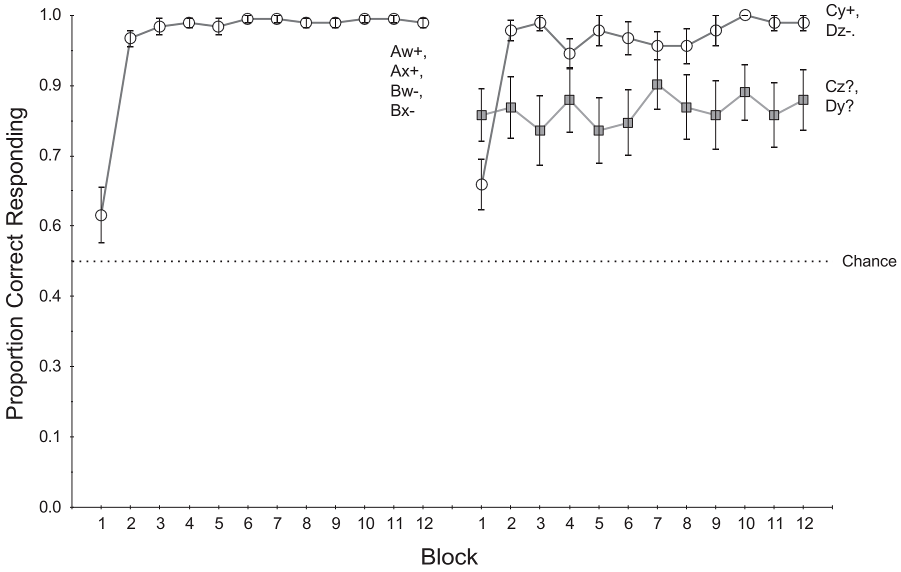
We were particularly interested in the relationship of these biases in looking toward the most predictive elements of compounds on subsequent test performance (cf. Le Pelley et al., 2011). The time spent during Stage-1 training looking at the irrelevant dimensions' ROIs was subtracted from those of the relevant dimension, obtaining a single datum that reflects each participant's bias. This correlated, reliably, with the overall accuracy, $r(30) = .41$, $p < .001$, in the test that followed Stage 1. The correlation retained reliability when Stage-1 acquired equivalence and optional-shift discrimination were used to control for nonspecific variation, $r(27) = .40$, $p < .014$. This was achieved by averaging the performance of all Stage-1 trials of both tasks. These findings also match those of Le Pelley et al.'s, which they interpreted as evidence that highly predictive stimuli, provoke overt attention and, therefore, benefit from efficient learning.

General Discussion

The two experiments reported here found that test performance in configural acquired equivalence (e.g., Coutureau et al., 2002; Honey & Ward-Robinson, 2001, 2002; Honey & Watt, 1998; Iordanova et al., 2007; Robinson & Owens, 2013; Ward-Robinson & Honey, 2000; see also Delamater & Joseph, 2000) and optional shift (e.g., Duffaud et al., 2007; see also, Kendler et al., 1964; Schwartz et al., 1971) was positively correlated. We noted that that these correlations could have been generated by nonspecific variables such as motivation or arousal. However, we found that the correlations retained their reliability when Stage-1 performance was included as a control variable. Both procedures were similar in that they required learning about outcomes that were signaled by cues that were presented in compounds; and similar in that learning occurred over multiple stages before testing. On that basis, and for reasons of parsimony, we sought to test the

Figure 5

Means and Their Standard Errors From the Optional-Shift Task From Experiment 2



Note. Performance is expressed as the proportion of correct responses, with 0.5 representing chance which is represented by a horizontal, dashed line. A–D, w–z, +/− and ? represent the format of the trials in the experiment. A–D and w–z represent experimental stimuli. + and − represent outcomes that participants predicted from A–D and w–z. ? indicates that no outcome was presented. The leftmost panel summarizes data from Stage 1, the initial Aw+, Ax+, Bw−, and Bx− discriminations. The rightmost panel summarizes data from the transfer trials (A+, B−, circles) and test trials (C?, D?, squares) in Stage 2. Stage 1 and Stage 2 data are blocked over, respectively, eight and three trials.

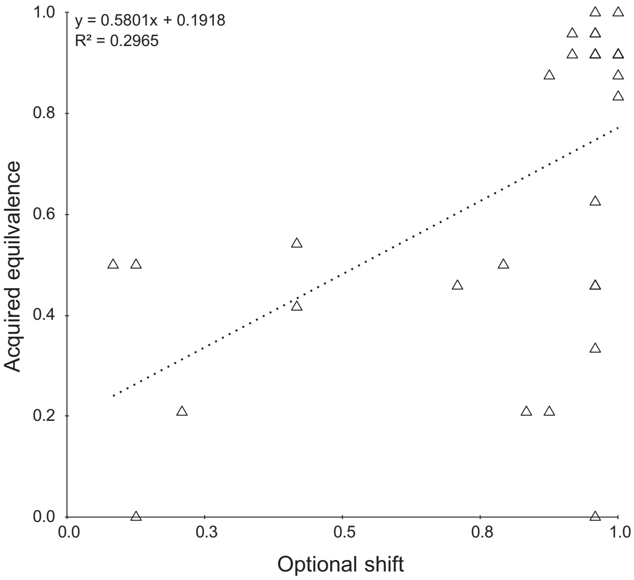
possibility that a common psychological mechanism was responsible for learning in both tasks, and this was supported by our new finding that test performance in the two tasks was positively correlated. This finding, alone, does not point to any specific model or mechanism but it can be understood, in general terms, by assuming that a common, underpinning mechanism will vary in its efficiency across participants and that this will be reflected in the performance on both tasks.

From here, we consider if any extant models of learning could explain both configural acquired equivalence and optional shift, found here, which is a requirement of applying a common mechanism to the two phenomena. Duffaud et al. (2007) reason that Mackintosh's (1975) model anticipates optional shift. During Stage 1 (Aw+, Ax+, Bw−, and Bx−) A and B's outcomes are entirely reliable, which will increase their associability. As a consequence of this change, more will be learned about these stimuli than their partners (y and z) in the Stage-2 transfer trials (Cy+ and Dz−). Thus, the test responding to Cz? and Dy? will reflect the transfer-trial outcomes from C and D, rather than, y and z. However, it is unclear how such changes in associability could explain configural acquired equivalence. Honey and Ward-Robinson (2002) also point out that configural acquired equivalence is beyond the scope of Pearce's (1994) model of learning. It correctly predicts the solution of the configural, Stage-1 discrimination learning (Aw+, Ax−, Bw−, Bx+, Cw+, Cx−, Dw−, and Dx+), where configural units represent each trial type; but there is no means to explain the final test result following the transfer trials. Furthermore, Pearce's model appears unable to accommodate the optional-shift effect because, for example, there would be no reason to expect more response transfer from Cy+ to Cz? than to Dy?

However, two other models may be able to provide an account of these results. Delamater (2012) reported accurate simulations of several discrimination phenomena and his model may be considered as a possible explanation of the phenomena that we present here. One of the discriminations that Delamater simulates is one of his demonstrations of acquired equivalence and distinctiveness (Delamater, 1998). In that experiment, rats received discriminations with a pair of auditory stimuli (A1 and A2) and a pair of visual stimuli (V1 and V2) which were either reinforced by delivery of food or sucrose (+ and *), or were not reinforced (−). The initial discrimination took the form: A1+, A2−, V1−, and V2*. This discrimination was reversed in secondary discrimination, where one group of rats, whose reinforcing outcomes were signaled by the stimulus from the original modality (i.e., A1−, A2+, V1*, and V2−), learned more slowly than a second group whose reinforcing outcomes were used in the alternative modality (i.e., A1−, A2*, V1+, and V2−). Delamater's model is a three-layer network whose connection strengths are modified by standard back-propagation rules. The network's input and output layers are static and correspond, respectively, to the to-be-discriminated stimuli and their outcomes. However, discrimination training gradually shapes the selection of the hidden layer's units and it is this feature that provides the model's successes. In particular, these units will converge or be differentiated when stimuli and their outcomes are, respectively, relatively similar or relatively dissimilar. We can conceive of pairs of stimuli being relatively similar within

Figure 6

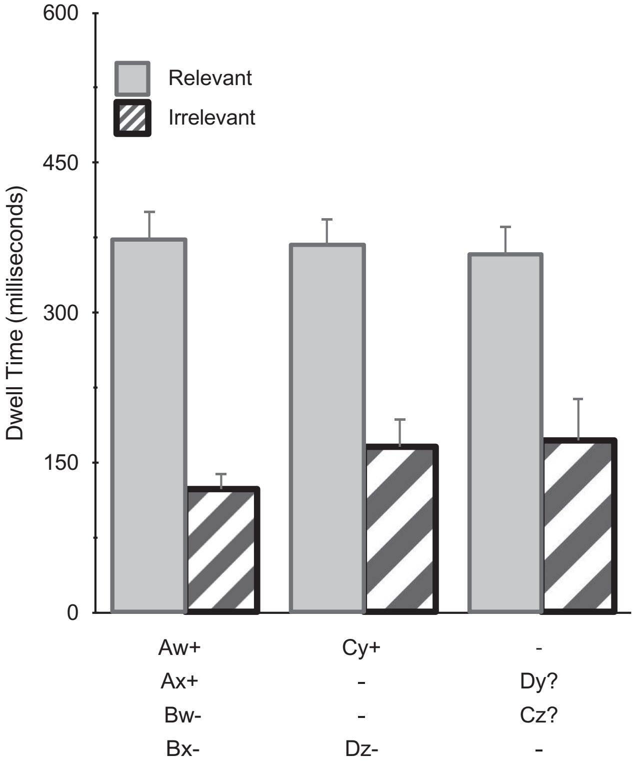
Scattergram of the Relationships Between Performance on the Configural Acquired Equivalence (Ordinate) and Optional-Shift Task (Abscissa) in Experiment 2



Note. Performance is expressed as the proportion of correct responses, with 0.5 representing chance. Lines represent the linear regression of optional-shift performance on the configural acquired equivalence task, whose values are presented in the top left of the figure with R^2, the proportion of variance in performance in acquired equivalence predicted by optional-shift performance.

Figure 7

Means and Their Standard Errors From the Eye-Tracking Data of the Experiment 2 Optional-Shift Task



Note. These eye-tracking data were taken concurrently with the performance data, summarized in Figures 4, 5 and 7. A–D, w–z, +/− and ? represent the format of the trials in the experiment. A–D and w–z represent experimental stimuli. + and − represent outcomes that participants predicted from A–D and w–z. ? indicates that no outcome was presented. The leftmost panel summarizes data from Stage 1, the initial Aw+, Ax+, Bw−, and Bx− discriminations. The central panel summarizes data from the transfer in Stage 2 (Cy+, Dz−), and the rightmost panel summarizes data from the test in Stage 2 (Cz?, Dy?). In each panel, data are separately presented for the elements of each compound trial: Stimuli A–D are relevant and w–z are irrelevant to the discriminations.

their modality and relatively dissimilar between their modalities (e.g., A1 is more similar to A2 than it is to V2). And although reinforcement (+ or *) can be conceived of as dissimilar to nonreinforcement, each reinforcer is discriminable from the other, that is, dissimilar.

Put simply, the model accounts for Delamater's (1998) between-group difference in reversal learning because, although different reinforcers will generally aid discrimination, they will be most helpful when discriminating relatively similar stimuli—here those from the same modality. More specifically, the model anticipates that the original discrimination from Delamater (1998; A1+, A2−, V1−, and V2*) will result in sharing of hidden units within each modality, based on their relative similarities to each other. That is, there will be a tendency for within-modality generalization that will restrict the development of discrimination. However, this will be more than offset by hidden-layer differentiation from the differential reinforcement, within each modality. In the second discrimination, the reinforced and nonreinforced outcomes will maintain hidden-layer differentiation within each modality. But this differentiation will be most marked in A1−, A2*, V1+, and V2− because, within each modality, each stimulus pair is reinforced with a different reinforcer than it was in the original discrimination. The reverse is true for the slower reversal, A1−, A2+, V1*, and V2−, where the same reinforcer is used across the original and the reversed discriminations, within each modality.

Delamater (2012) does not provide a simulation of the configural acquired equivalence procedure that we report here, but his model is similar to the model described by Honey and colleagues (e.g., Honey & Ward-Robinson, 2002; Honey et al., 2010), which was designed to account for configural acquired equivalence. Rather than employing back-propagation, Honey and Ward-Robinson, describe a three-layer Hebbian network, whose hidden layer, is shaped by similarity and trial outcomes. That is, although the learning rules of Delamater and Honey and Ward-Robinson differ, both obtain comparable modifications of their hidden layers to accurately explain discrimination. For this reason, it seems likely that both of these models would accommodate the configural acquired equivalence results reported here. Robinson et al. (2019) have simulated the Honey and Ward-Robinson model and confirmed that it will correctly anticipate the results of several configural forms of acquired equivalence. Furthermore, Bru García (2021, pp. 231–236) found that it can also capture the results of Delamater's (1998) nonconfigural acquired equivalence experiment, summarized above. Thus, it seems that both models would provide adequate accounts of the configural acquired equivalence. To fully account for the correlations found in this report, it would be necessary for a single model to account for optional shift, as well as configural acquired equivalence. Neither of the two models described, here, has attempted to accommodate optional shift. However, Honey et al. (2010) have described how the Hebbian model can account for the intra-/extradimensional shift (e.g., George & Pearce, 1999; Mackintosh & Little, 1969), which, like the optional shift, involves learning about the relevance of different dimensions to their outcomes across multiple stages (see, e.g., Duffaud et al., 2007 for discussion of their similarities).

Experiment 2 also measured selective looking at elements of the compound stimuli. In all stages of the optional-shift procedure, participants spent longer looking at the element of each compound that reliably predicted its outcome, over the alternative element that was not predictive. The bias seen in Stage 1 of optional shift positively correlated with the subsequent test performance. Both of these findings have been reported previously by Le Pelley et al. (2011; see also, Beesley & Le Pelley, 2011; Griffiths & Le Pelley, 2019). In explaining these findings in their own results, Le Pelley et al. (2011) consider models of learning that include attentional components, driven by a cue's predictiveness. They draw on the suggestions of Mackintosh (1975; see also Kruschke, 1996) that predictive cues will retain high associability, which may be a proxy for attention; and, in turn, this will support superior learning. However, these authors acknowledge that, although this pattern of results is consistent with the idea that learning is driven by attention, which is driven by predictability, it does so only through correlation. Thus it is possible, instead, that attention and future performance are, themselves, causally independent but are both positively affected by the predictability of cues. And more generally, it should be acknowledged that gaze toward visual stimuli, as measured in eye-tracking experiments, may not measure attention.

Our evidence for a common mechanism in configural acquired equivalence and optional shift has other parallels in existing reports, where performance-influencing variables are explicitly manipulated. Robinson and Owens' (2013) found that healthy elderly participants' configural acquired equivalence was reduced relative to a younger group. And aging has been reported to reduce the attentional shift in people (e.g., Robbins et al., 1998; Sahakian et al., 1990) and in rodents (Barense et al., 2002), albeit in an intra-/extradimensional shift (e.g., George & Pearce, 1999; Mackintosh & Little, 1969), rather than in optional shift. Additional parallels come from the effects of brain lesions. Patients with frontal damage have been reported to have reduced capacity on intra-/extradimensional shift performance (Owen et al., 1991). Matching findings have been reported in nonhuman subjects using experimental lesions of the basolateral prefrontal cortex in primates (e.g., Dias et al., 1997; Roberts et al., 1994) and mediolateral prefrontal cortex in rats (Birrell & Brown, 2000; for a review see, George et al., 2010). The mediolateral prefrontal cortex has also been shown to be involved in configural acquired equivalence in rats (Iordanova et al., 2007). A second, though less often implicated, brain region with such parallel effects is the entorhinal cortex: Selective lesions in rats have been found to modify performance on both configural acquired equivalence (Coutureau et al., 2002) and intra-/extradimensional shift (Oswald et al., 2001), which used similar tasks and stimuli. Our current findings add to this evidence and have the additional strength of comparing task performance within the same experiment. This strategy has allowed us to examine potential correlations between participants' performance, which would otherwise have been impossible.

References

Audacity. (Version 2.3) [Computer software]. https://audacityteam.org

Barense, M. D., Fox, M. T., & Baxter, M. G. (2002). Aged rats are impaired on an attentional set-shifting task sensitive to medial frontal cortex damage in young rats. *Learning & Memory*, *9*(4), 191–201. https://doi.org/10.1101/lm.48602

Beesley, T., & Le Pelley, M. E. (2011). The influence of blocking on overt attention and associability in human learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(1), 114–120. https://doi.org/10.1037/a0019526

Birrell, J. M., & Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *The Journal of Neuroscience*, *20*(11), 4320–4324. https://doi.org/10.1523/JNEUROSCI.20-11-04320.2000

Bru García, S. (2021). *Testing a connectionist model of acquired equivalence* [Doctoral thesis, University of Nottingham]. https://eprints.nottingham.ac.uk/64272/

Coutureau, E., Killcross, A. S., Good, M., Marshall, V. J., Ward-Robinson, J., & Honey, R. C. (2002). Acquired equivalence and distinctiveness of cues: II. Neural manipulations and their implications. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(4), 388–396. https://doi.org/10.1037/0097-7403.28.4.388

Delamater, A. R. (1998). Associative mediational processes in the acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*(4), 467–482. https://doi.org/10.1037/0097-7403.24.4.467

Delamater, A. R. (2012). On the nature of CS and US representations in pavlovian learning. *Learning & Behavior*, *40*(1), 1–23. https://doi.org/10.3758/s13420-011-0036-4

Delamater, A. R., & Joseph, P. (2000). Common coding in symbolic matching tasks with humans: Training with a common consequence or antecedent. *The Quarterly Journal of Experimental Psychology Section B*, *53*(3b), 255–273. https://doi.org/10.1080/713932729

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin card sort test: Restriction to novel situations and independence from "on-line" processing. *The Journal of Neuroscience*, *17*(23), 9285–9297. https://doi.org/10.1523/JNEUROSCI.17-23-09285.1997

Duffaud, A. M., Killcross, A. S., & George, D. N. (2007). Optional-shift behaviour in rats: A novel procedure for assessing attentional processes in discrimination learning. *Quarterly Journal of Experimental Psychology*, *60*(4), 534–542. https://doi.org/10.1080/17470210601154487

George, D. N., Duffaud, A. M., & Killcross, S. (2010). Neural correlates of attentional set. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 351–383). Oxford University Press.

George, D. N., & Pearce, J. M. (1999). Acquired distinctiveness is controlled by stimulus relevance not correlation with reward. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*(3), 363–373. https://doi.org/10.1037/0097-7403.25.3.363

Griffiths, O., & Le Pelley, M. E. (2019). The outcome predictability bias is evident in overt attention. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(3), 290–300. https://doi.org/10.1037/xan0000210

Honey, R. C., Close, J., & Lin, T. C. E. (2010). Acquired distinctiveness and equivalence: A synthesis. In C. Mitchell & M. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 159–186). Oxford University Press.

Honey, R. C., & Ward-Robinson, J. (2001). Transfer between contextual conditional discriminations: An examination of how stimulus conjunctions are represented. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*(3), 196–205. https://doi.org/10.1037/0097-7403.27.3.196

Honey, R. C., & Ward-Robinson, J. (2002). Acquired equivalence and distinctiveness of cues: I. Exploring a neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(4), 378–387. https://doi.org/10.1037/0097-7403.28.4.378

Honey, R. C., & Watt, A. (1998). Acquired relational equivalence: Implications for the nature of associative structures. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*(3), 325–334. https://doi.org/10.1037/0097-7403.24.3.325

Iordanova, M. D., Killcross, A. S., & Honey, R. C. (2007). Role of the medial prefrontal cortex in acquired distinctiveness and equivalence of cues. *Behavioral Neuroscience*, *121*(6), 1431–1436. https://doi.org/10.1037/0735-7044.121.6.1431

James, W. (1890). *The principles of psychology*. Holt.

Jamovi. (Version 2.3.28.0) [Computer software]. https://www.jamovi.org

Kelley, K. (2007). Confidence intervals for standardized effect sizes: Theory, application, and implementation. *Journal of Statistical Software*, *20*(8), 1–24. https://doi.org/10.18637/jss.v020.i08

Kendler, T. S., Kendler, H. H., & Silfen, C. K. (1964). Optional shift behavior of albino rats. *Psychonomic Science*, *1*(1–12), 5–6. https://doi.org/10.3758/BF03342760

Kim, S. (2015). Ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, *22*(6), 665–674. https://doi.org/10.5351/CSAM.2015.22.6.665

Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 3–26. https://doi.org/10.1037/0278-7393.22.1.3

Le Pelley, M. E., Beesley, T., & Griffiths, O. (2011). Overt attention and predictiveness in human contingency learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(2), 220–229. https://doi.org/10.1037/a0021384

Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *The Quarterly Journal of Experimental Psychology Section B*, *56*(1b), 68–79. https://doi.org/10.1080/02724990244000179

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276–298. https://doi.org/10.1037/h0076778

Mackintosh, N. J., & Little, L. (1969). Intradimensional and extradimensional shift learning by pigeons. *Psychonomic Science*, *14*(1), 5–6. https://doi.org/10.3758/BF03336395

MBESS (Version 4.9.3.) [Computer software]. https://cran.r-project.org/web/packages/MBESS/index.html

Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. Yale University Press.

Oswald, C. J. P., Yee, B. K., Rawlins, J. N. P., Bannerman, D. B., Good, M., & Honey, R. C. (2001). Involvement of the entorhinal cortex in a process of attentional modulation: Evidence from a novel variant of an IDS/EDS procedure. *Behavioral Neuroscience*, *115*(4), 841–849. https://doi.org/10.1037/0735-7044.115.4.841

Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., & Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shirting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, *29*(10), 993–1006. https://doi.org/10.1016/0028-3932(91)90063-E

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, *101*(4), 587–607. https://doi.org/10.1037/0033-295X.101.4.587

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Robbins, T., James, M., Owen, A., Sahakian, B., Lawrence, A., Mcinnes, L., & Rabbitt, P. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society*, *4*(05), 474–490. https://doi.org/10.1017/S1355617798455073

Roberts, A. C., De Salvia, M. A., Wilkinson, L. S., Collins, P., Muir, J. L., Everitt, B. J., & Robbins, T. W. (1994). 6-Hydroxydopamine Lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin card sort test: Possible interactions with subcortical dopamine. *The Journal of Neuroscience*, *14*(5), 2531–2544. https://doi.org/10.1523/JNEUROSCI.14-05-02531.1994

Robinson, J., George, D. N., & Heinke, D. (2019). A computational implementation of a Hebbian learning network and its application to configural forms of acquired equivalence. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(3), 356–371. https://doi.org/10.1037/xan0000203

Robinson, J., & Owens, E. (2013). Diminished acquired equivalence yet good discrimination performance in older participants. *Frontiers in Psychology*, *4*, Article 726. https://doi.org/10.3389/fpsyg.2013.00726

RStudio (Version 2023.06.1+524) [Computer software]. https://posit.co/products/open-source/rstudio/

Sahakian, B. J., Downes, J. J., Eagger, S., Everden, J. L., Levy, R., Philpot, M. P., Roberts, A. C., & Robbins, T. W. (1990). Sparing of attentional relative to mnemonic function in a subgroup of patients with dementia of the Alzheimer type. *Neuropsychologia*, *28*(11), 1197–1213. https://doi.org/10.1016/0028-3932(90)90055-S

Schwartz, R. M., Schwartz, M., & Tees, R. C. (1971). Optimal intradimensional and extradimensional shifts in the rat. *Journal of Comparative and Physiological Psychology*, *77*(3), 470–475. https://doi.org/10.1037/h0031886

van Aert, R. C. M., & Goos, C. (2023). A critical reflection on computing the sampling variance of the partial correlation coefficient. *Research Synthesis Methods*, *14*(3), 520–525. https://doi.org/10.1002/jrsm.1632

Ward-Robinson, J., & Honey, R. C. (2000). A novel contextual dimension for use with an operant chamber: From simple to hierarchical forms of learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*(3), 358–363. https://doi.org/10.1037//0097-7403.26.3.358

Weatherburn, C. E. (1961). *A first course in mathematical statistics*. The English Language Book Society And Cambridge University Press. Retrieved May 6, 2024, from https://archive.org/details/firstcourseinmat029137mbp/page/n273/mode/2up

Yule, G. U. (1919). *An introduction to the theory of statistics* (5th ed.). Charles Griffin and Company. Retrieved April 24, 2024, from https://archive.org/details/anintroductiont00yulegoog/page/n266/mode/2up

Received May 16, 2021
Revision received May 24, 2024
Accepted June 12, 2024 ■