Learn Behav (2012) 40:530–541
DOI 10.3758/s13420-012-0072-8

Contextual modulation of attention in human category learning

David N. George · John K. Kruschke

Published online: 15 April 2012

Abstract In a category-learning experiment, we assessed whether participants were able to selectively attend to different components of a compound stimulus in two distinct contexts. The participants were presented with stimulus compounds for which they had to learn categorical labels. Each compound comprised one feature from each of two dimensions, and on different trials the compound was presented in two contexts, X and Y. In Context X, Dimension A was relevant to the solution of the categorization task and Dimension B was irrelevant, whereas in Context Y, Dimension A was irrelevant and Dimension B was relevant. The results of transfer tests to novel stimuli suggested that people learned to attend selectively to Dimension A in Context X and Dimension B in Context Y. These findings contribute to the growing body of evidence that people can learn to selectively attend to particular dimensions of stimuli dependent on the context in which the stimuli are presented. Furthermore, the findings demonstrate that context-dependent changes in attention transfer to other categorization tasks involving novel stimuli.

Keywords Attention · Categorization · Context · Human learning · Associative learning

D. N. George (✉)
Department of Psychology, University of Hull,
Hull, UK
e-mail: d.george@hull.ac.uk

D. N. George
School of Psychology, University of New South Wales,
Sydney, New South Wales, Australia

J. K. Kruschke
Department of Psychological and Brain Sciences,
Indiana University,
Bloomington, IN, USA

People and animals may profit by attending to different aspects of their environment in different situations. For example, while driving an automobile, people should attend to speed limit signs, yet while walking beside the same road, people may ignore such signs. While playing football, jersey color is highly relevant, but while walking through a crowded lobby, people may ignore clothing color. A number of laboratory-based studies investigating context-dependent partitioning of knowledge about category boundaries have shown that people can learn to utilize different types of information in different contexts (e.g., Aha & Goldstone, 1992; Sewell & Lewandowsky, 2011; Yang & Lewandowsky, 2003). The transfer of such learning to new categorization tasks has not, however, been explored in any depth, and many theories of attentional learning have neglected the role of context. In this article, we report clear evidence that context-specific learned attention can transfer to categorization tasks involving novel stimuli, and we suggest extensions of a formal model that might address this phenomenon.

The intradimensional–extradimensional (ID–ED) shift effect has been described as providing "perhaps the best evidence that transfer between discrimination problems may be based partly on increases in attention to relevant dimensions and decreases in attention to irrelevant dimensions" (Mackintosh, 1974, p. 597). In one demonstration of the effect, George and Pearce (1999, Exp. 1; see also Kruschke, 1996; Mackintosh & Little, 1969) trained pigeons on a discrimination task in which the stimuli were colored lines of different orientations. For some pigeons, the color of the lines signaled whether or not food would be made available, but the orientation of the lines was irrelevant. For other birds, the color was irrelevant and the orientation signaled which lines were paired with food. Once they had mastered this discrimination, the pigeons learned a second discrimination more rapidly when it was based on the

previously relevant property of the stimuli than when it was based on the previously irrelevant property of the stimuli.

In many models of associative learning, the attention paid to a stimulus (or its *associability*) may change as a result of experience (e.g., Kruschke, 1992, 2001; Mackintosh, 1975; Pearce, George, & Redhead, 1998; Pearce & Hall, 1980), but these models tend not to allow different amounts of attention to be paid to the same stimulus presented in different contexts. Kruschke (2009), however, has argued that some aspects of the inverse base-rate effect are best explained by a model that allows for the amount of attention paid to a stimulus to be modulated by the context in which that stimulus is presented. Although Griffiths and Le Pelley (2009) found no evidence of context-dependent attention in a series of experiments employing a blocking design, Sloutsky and Fisher (2008) showed that young children were capable of flexibly responding to a compound stimulus on the basis of different features in distinct contexts. One possible reason why Griffiths and Le Pelley may have found no effect of context on attention will be considered in the Discussion section.

Sloutsky and Fisher (2008) trained 4- to 5-year-old children on a discrimination task in which they were presented with triplets of colored shapes in each of two contexts (defined by the location of the triplet on the computer screen and the color of a background rectangle). Each triplet comprised a single target item and two test items. The participants were rewarded for selecting one of the test items, but not the other. In Context X, all three items were the same color, and shape was predictive of the correct test item. In Context Y, however, all items were the same shape, and color was predictive. Following training, test trials were administered with ambiguous test triads. For these triads, one of the test items differed from the target in color but not in shape, whereas the other test item differed from the target in shape but not in color. Sloutsky and Fisher found that a higher proportion of shape-based choices than color-based choices were made to the testing triads in Context X, and the reverse pattern was observed in Context Y.

Sloutsky and Fisher (2008) concluded that the type of flexible behavior described above could be explained in terms of relatively simple mechanisms of associative and attentional learning. Although their experiment was undeniably elegant and ingenious, it did not address the question of whether any such changes in attention may transfer to new learning situations, and it did not provide unambiguous evidence that people can learn to selectively attend to different features of a stimulus in different contexts. Rather, it is still possible that responses to the test triads simply reflected similarities between the training and test triads. We have conducted a number of computer simulations of two models of associative learning (Pearce, 1994; Rescorla & Wagner, 1972) based on Sloutsky and Fisher's experimental design, and we found that they predict the pattern of responding observed by Sloutsky and Fisher without recourse to mechanisms of attentional change (see the Appendix).

The purpose of the experiment reported here was to determine whether the types of attentional effects reported by George and Pearce (1999) can be modulated by the context in which cues are presented. We employed a multiple-learning-phase design in which different stimuli (differing along the same two dimensions) were used in different learning phases. The advantage of using this type of design is that, with appropriate counterbalancing of the stimuli, primary generalization between training and test patterns should not lead to any systematic differences in the associative strengths of individual features of the test patterns, in the absence of changes in attention. As a consequence, if we were to observe an effect consistent with context-dependent attention, it would not be predicted by models of associative learning that do not include an attentional component.

Our participants were trained on a variant of an optional-shift design (e.g., Duffaud, Killcross, & George, 2007; Kendler & Kendler, 1964). A standard optional-shift experiment consists of two stages of training. In the first stage, participants are trained on a discrimination task in which stimulus compounds differ on two dimensions, but one dimension is relevant to the solution of the discrimination, and the second dimension is irrelevant (i.e., A1B1→C1, A1B2→C1, A2B1→C2, and A2B2→C2, where A1 and A2 are features belonging to Dimension A, B1 and B2 are features belonging to Dimension B, and C1 and C2 are different outcomes). In a second stage of training, a new discrimination is learned between stimulus compounds comprising novel exemplars from each dimension. In this stage, however, only compounds differing on both dimensions are trained (i.e., A3B3→C1 and A4B4→C2). Finally, test trials are given with the other combinations of the features trained in Stage 2 (A3B4 and A4B3). If the initial training causes greater attention to the relevant dimension (A) than to the irrelevant dimension (B), these test compounds should be perceived as more similar to the training compounds with which they share a feature on Dimension A than to those with which they share a feature on Dimension B. That is, participants should judge A3B4 to be more similar to A3B3 than it is to A4B4 and, hence, should rate C1 as a more likely category than C2 in the presence of A3B4. Our new design incorporated copies of the optional-shift design in two different contexts, such that different dimensions were relevant in different contexts.

Table 1 shows the design of our experiment, along with an example of the training received by a participant. The participants were trained on a series of concurrent categorization tasks in two contexts made distinct from each other by varying the background color of the computer screen (red or green). The stages of training are indicated in the left column of Table 1. The first two data columns of Table 1

Table 1 The design of the experiment

	Abstract Design		Particular Instantiation	
	Context X	Context Y	Red	Green
Stage 1	**A1**B1 → C1	A1**B1** → C3	**Raccoon**–Handsaw → F	Raccoon–**Handsaw** → H
	A1B2 → C1	A1**B2** → C4	**Raccoon**–Wrench → F	Raccoon–**Wrench** → J
	A2B1 → C2	A2**B1** → C3	**Squirrel**–Handsaw → G	Squirrel–**Handsaw** → H
	A2B2 → C2	A2**B2** → C4	**Squirrel**–Wrench → G	Squirrel–**Wrench** → J
Stage 2	A3B3 → C1	A5B5 → C3	Beaver–Pliers → F	Skunk–Hammer→ H
	A3B4 → ?	A5B6 → ?	Beaver–Drill → ?	Skunk–Screwdriver → ?
	A4B3 → ?	A6B5 → ?	Chipmunk–Pliers → ?	Rabbit–Hammer→ ?
	A4B4 → C2	A6B6 → C4	Chipmunk– Drill → G	Rabbit–Screwdriver → J
Stage 3	**A3**B3 → C1	**A5**B5 → C3	**Beaver**–Pliers → F	**Skunk**–Hammer→ H
	A3B4 → C1	**A5**B6 → C3	**Beaver**–Drill → F	**Skunk**–Screwdriver → H
	A4B3 → C2	**A6**B5 → C4	**Chipmunk**–Pliers → G	**Rabbit**–Hammer→ J
	A4B4 → C2	**A6**B6 → C4	**Chipmunk**–Drill → G	**Rabbit**–Screwdriver → J

A1 to A6 and B1 to B6 represent the six items on each stimulus dimension (animals and hand tools). The relevant dimension for each stimulus is shown in boldface. C1 to C4 represent the four categories (F, G, H, and J). Contexts X and Y were the background colors red and green. The assignment of values to each of these identifiers was randomized for each participant. Within each block of eight trials in a stage, each of the eight trial types was presented once. Examples of the stimuli presented to a participant in each stage and their category membership are also show.

show the abstract design, indicating the dimensions and outcomes using abstract symbols. The last two columns of Table 1 show a particular instantiation of the abstract dimensions (e.g., abstract feature A1 is instantiated as the word "raccoon"). The assignment of specific instantiations to abstract values was randomized across participants.

The column of Table 1 labeled "Context X" lists the training trials to which a participant was exposed in one of the contexts. The first two stages of training in this context constituted an optional-shift design of the type described above (with the modification that the test trials with compounds A3B4→? and A4B3→? were intermixed with continued training on A3B3→C1 and A4B4→C2). Notice that Dimension A is relevant in Stage 1 for Context X. The column of Table 1 labeled "Context Y" shows the same type of optional-shift design, but with Dimension B relevant instead of Dimension A.

Considering the particular instantiation shown in Table 1, for the red context during Stage 1, the participant learned that the word pairs raccoon–handsaw and raccoon–wrench belonged to category F, and the word pairs squirrel–handsaw and squirrel–wrench belonged to category G. Hence, "raccoon" and "squirrel" signaled category membership (shown in bold typeface in Table 1), whereas "handsaw" and "wrench" provided no information about category membership. Thus, the dimension "animals" was relevant and the dimension "hand tools" was irrelevant.

During the second stage of the experiment, novel values were presented from each dimension, but feedback was only given on some trials. For example, the participant learned that beaver–pliers and chipmunk–drill belonged to categories F and G, respectively. These Stage 2 items could be learned by attending to either dimension (or both), and therefore selective attention by the participant was optional. The remaining trials given in Stage 2 were designed to determine whether participants did selectively attend to one of the dimensions (and whether it was the same dimension as the relevant one from Stage 1 in each context). On these test trials, the other combinations of the two animals and two hand tools (beaver–drill and chipmunk–pliers) were presented, and participants were asked to categorize them, but no corrective feedback was given. If participants learned the context-dependent relevant dimensions in Stage 1 and continued to pay attention to those dimensions when learning Stage 2, the responses on the test trials should reflect dimension-based generalization. For example, if the participants paid attention to the animal dimension when learning beaver–pliers→F and chipmunk–drill→G, the test item beaver–drill should be responded to with F, not G. On the other hand, if Stage 1 training had no influence on attention to the two stimulus dimensions, the participants should be equally likely to respond with F or G.

Inspection of the final column of Table 1 reveals that the participants received different training in the green context. Here, hand tools were relevant and animals were irrelevant during Stage 1. If attention may be modulated by context, then at the end of Stage 1, participants should attend more to animals than to hand tools in the red context, but more to hand tools than to animals in the green context. Hence, categorization responses on the probe test trials with the word pairs

skunk–screwdriver and rabbit–hammer in the green context should reflect greater similarity of word pairs containing the same hand tool than of those containing the same animal.

Stage 3 of our design was another new extension of the optional-shift paradigm. In Stage 3, all stimuli had novel values on the dimensions. The corrective feedback suggested that the same dimension remained relevant in Context X, but that the other dimension became relevant in Context Y. Thus, Stage 3 constituted an intradimensional shift in Context X, but an extradimensional shift in Context Y. The purpose of including this stage was to determine whether any effect of the contextual modulation of attention observed in Stage 2 was resistant to explicit changes in feedback.

For example, inspection of the bottom row of Table 1 reveals that animals were now relevant to the solution of the categorization task, and hand tools were irrelevant, in *both* the red and the green contexts. Here, contextual modulation of attention was expected to result in more rapid acquisition of the categorization task learned in the red context, in which animals had been relevant during Stage 1 (intradimensional shift), than of the task presented in the green context, in which animals had previously been irrelevant (extradimensional shift).

Method

Participants

A group of 78 undergraduate students of psychology at Indiana University, Bloomington, were given credit toward course requirements for their participation.

Although all participants received the same type of training across the three stages of the experiment, they were divided into three groups that received slightly different treatments with respect to the trial sequencing and the duration of the contextual cues (details given below). The reason for these differences is beyond the scope of this article, and the three groups were statistically equivalent. Consequently, all analyses of the data were collapsed across the three treatment groups. The data from a further 23 participants were excluded because they failed to achieve the learning criterion (explained below).

Apparatus and stimuli

An IBM-compatible PC running the Windows NT 4 operating system and programmed using Visual Basic 6 (Microsoft Corp., Redmond, WA) controlled the presentation of the stimuli on a 38-cm color monitor and recorded responses on a standard keyboard. The stimuli were the names of six animals indigenous to North America ("raccoon," "squirrel," "beaver," "chipmunk," "skunk," and "rabbit") and of six common hand tools ("handsaw," "wrench," "pliers," "drill," "hammer," and "screwdriver") printed in black letters 1.25 cm high in Times New Roman typeface. The context was provided by background illumination of the entire screen in either red or green. At all other times, the background illumination of the screen was a mid-gray. On training trials, the stimulus words were presented 10 cm from the top of the screen, and the distance between the centers of the two words was 15 cm.

For half of the participants, the category "animals" served as Dimension A, and the category "hand tools" served as Dimension B. For the remaining participants in each group the assignments of categories to Dimensions A and B were "hand tools" and "animals," respectively. For half of the participants in each of the groups, Context X was the red context and Context Y was the green context. For the other half of the participants, Context Y was the red context and Context X was the green context. For each participant, the six items in each category were randomly assigned to the six values on the appropriate dimension (A1–A6 and B1–B6) at the beginning of the experimental session.

Procedure

The design of the experiment is shown in Table 1. On each trial, the context, one item from Dimension A, and one item from Dimension B were presented, and the participant was asked which of four categories (F, G, H, or J) the word pair belonged to. The participant could respond by pressing one of four keys labeled "F," "G," "H," and "J" on the computer keyboard. Assignment of these responses to Categories C1–C4 was randomly determined for each participant prior to the start of the experiment. After a response was made, the stimuli remained on the screen, and feedback (the word "Correct" or "Incorrect") appeared. The correct answer was then supplied. If the participant took longer than 30 s of study time, a warning appeared, and the next trial started automatically. Between trials, the screen was blank for approximately 2 s. On those trials in which no feedback was given during Stage 2 (indicated by a question mark in Table 1), the following message was displayed "No information is currently available."

During each stage of the experiment, each block of eight trials contained the eight cases shown in the appropriate row of Table 1, randomly permuted within certain constraints, as follows. For the first group of participants, the first four trials within a given block were all presented in Context X, whereas the last four trials were presented in Y. For these participants, the subsequent block of trials consisted of four trials in Context Y followed by four trials in Context X. In this way, excluding the first and last sets of trials of the experiment, the participants always experienced alternating "blocks" of eight consecutive trials in each context. The appropriate contextual cues were presented at the beginning of the intertrial interval (ITI) preceding a trial and remained on the screen until the end of the feedback event. Hence, one

of the contexts was present at all times during the experiment. The second group of participants received the same treatment, with the exception that the context was replaced by a blank gray screen during the ITI. For the final group of participants, the context was replaced by a blank gray screen during the ITI and no more than two successive trials were presented in the same context.

In Stages 1 and 3, training continued until the participant had achieved an accuracy of 87.5% correct (7 out of 8) on two successive blocks of eight trials. Since correct responses were available on only four of the trials in each block in Stage 2, this criterion was raised to 100% correct on two successive blocks of eight trials during this stage. If the accuracy criterion was not achieved after 400 trials within a given stage, the experiment was terminated and all data from that participant were discarded.

Results

A total of 23 participants failed to reach criterion during either Stage 1 or Stage 2, or else failed to complete the experiment within the allotted 1-h period. All of the data from these participants have been excluded from the analyses reported below. The performance of the remaining 78 participants in each of the three stages of the experiment is shown in Fig. 1. These survival curves suggest that each successive stage of the experiment was completed more rapidly than the previous stage.

Optional shift (Stage 2)

In order to yield a measure of the ability of each context to direct attention to the stimulus dimension that had been relevant in that context during Stage 1, we analyzed participants' responses on those trials for which feedback was not available during Stage 2. We separately calculated the proportion of responses to A3B4 and A4B3 that were consistent with generalization along Dimension A, and the proportion of responses to A5B6 and A6B5 that were consistent with generalization along Dimension B.

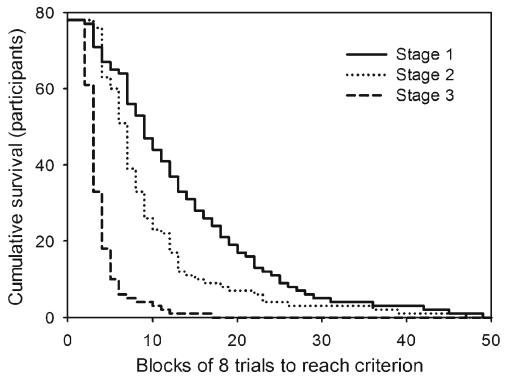

Fig. 1 Survival curves for each stage of the experiment

If, at the end of Stage 1, participants were attending more to Dimension A than to Dimension B in Context X, they should have treated A3B4 as though it was more similar to A3B3(→C1) than to A4B4(→C2). Therefore, the frequency of C1 responses to A3B4, denoted *f*(C1|A3B4), should be greater than the frequency of C2 responses to A3B4, denoted *f*(C2|A3B4). Likewise, *f*(C2|A4B3) should be greater than *f*(C1|A4B3). Stage 1 training should also have caused more attention to be paid to Dimension B than to Dimension A in Context Y. As a result, A5B6 should have been treated as if it were more similar to A6B6 than to A5B5, whereas A6B5 should have been treated more like A5B5 than A6B6. Hence, in each context we expected participants to respond to the test patterns according to the component of the word pair that belonged to the dimension that had been relevant in that context during Stage 1. Therefore, we defined an overall measure of a participant's tendency to respond according to Dimension A in Context X, denoted *O*(A|X), as follows:

$$O(\mathrm{A|X}) = \frac{f(\mathrm{C1|A3B4}) - f(\mathrm{C2|A3B4}) + f(\mathrm{C2|A4B3}) - f(\mathrm{C1|A4B3})}{f(\mathrm{C1|A3B4}) + f(\mathrm{C2|A3B4}) + f(\mathrm{C2|A4B3}) + f(\mathrm{C1|A4B3})}. \quad (1)$$

The magnitude of the overall attention to Dimension B in Context Y, *O*(B|Y), was calculated analogously:

$$O(\mathrm{B|Y}) = \frac{f(\mathrm{C4|A5B6}) - f(\mathrm{C3|A5B6}) + f(\mathrm{C3|A6B5}) - f(\mathrm{C4|A6B5})}{f(\mathrm{C4|A5B6}) + f(\mathrm{C3|A5B6}) + f(\mathrm{C3|A6B5}) + f(\mathrm{C4|A6B5})}. \quad (2)$$

Having calculated *O*(A|X) and *O*(B|Y), a measure of a participant's tendency to respond to the previously relevant stimulus dimension rather than the previously irrelevant stimulus dimension in each context was calculated by taking the arithmetic mean of these two numbers. This measure of optional-shift performance had a value in the range –1 to +1, where a score of 0 would indicate that the participant showed no systematic tendency to respond according to a particular dimension in each context, +1 would indicate that the participant's response was always governed purely by the element of a stimulus compound that belonged to the dimension previously relevant in each context, and a score of –1 would indicate that the participant always responded according to the previously irrelevant stimulus dimension in each context. Since one would expect participants to show no preference for either candidate category on trials without feedback until they had been exposed to the contingencies for the trained exemplars, responses from the first two blocks of trials in Stage 2 were excluded from this analysis.

The mean magnitude of the optional-shift effect shown by all 78 participants was .15. A one-sample Student's *t* test revealed that this value was significantly different from zero

[$t(77) = 2.96$, $p = .004$, $d = 0.34$]. Thus, in the optional-shift phase, people tended to pay more attention to the dimension that had previously been relevant in each specific context.

To provide some indication of the progression of the effect across the stage, Fig. 2 also shows the optional-shift effect calculated for the first two blocks of trials in Stage 2 and for the final two blocks of trials, in which each participant reached the learning criterion. Although the effect appears to be larger in the final two blocks than in the first two blocks, this difference was not reliable [$t(77) = 1.66$, $p = .1$].

Explicit shift (Stage 3)

Dimension A was relevant to the solution of the discrimination in both Contexts X and Y in this stage. Hence, participants received an intradimensional shift in Context X, in which Dimension A had previously been relevant during Stage 1, and an extradimensional shift in Context Y, in which Dimension A had previously been irrelevant and Dimension B had been relevant during Stage 1. Notice that because participants received corrective feedback on only half of the trials given in Stage 2, both stimulus dimensions were equally diagnostic in both contexts during that stage.

An explicit-shift score was computed for each participant in a manner similar to that for the optional-shift score. The degree of attention to Dimension A in Context X during the explicit-shift phase (that is, in the intradimensional shift) is denoted $I(\mathrm{A}|\mathrm{X})$. The formula for its calculation is shown in Eq. 3, below, and is identical to that in Eq. 1. The degree of attention to Dimension A in Context Y during the explicit-shift phase (i.e., in the extradimensional shift) is denoted $E(\mathrm{A}|\mathrm{Y})$. The formula for its calculation is shown in Eq. 4, below. This is like Eq. 2, except that the roles of C3 and C4 have been exchanged. In order to obtain a measure of any intradimensional shift advantage, we simply calculated the difference between these scores, $I(\mathrm{A}|\mathrm{X}) - E(\mathrm{A}|\mathrm{Y})$. A positive value of this explicit-shift score would indicate that the participant learned the intradimensional shift in Context X more rapidly than the extradimensional shift in Context Y.

$$I(\mathrm{A}|\mathrm{X}) = \frac{f(\mathrm{C1}|\mathrm{A3B4}) - f(\mathrm{C2}|\mathrm{A3B4}) + f(\mathrm{C2}|\mathrm{A4B3}) - f(\mathrm{C1}|\mathrm{A4B3})}{f(\mathrm{C1}|\mathrm{A3B4}) + f(\mathrm{C2}|\mathrm{A3B4}) + f(\mathrm{C2}|\mathrm{A4B3}) + f(\mathrm{C1}|\mathrm{A4B3})} \quad (3)$$

$$E(\mathrm{A}|\mathrm{Y}) = \frac{f(\mathrm{C3}|\mathrm{A5B6}) - f(\mathrm{C4}|\mathrm{A5B6}) + f(\mathrm{C4}|\mathrm{A6B5}) - f(\mathrm{C3}|\mathrm{A6B5})}{f(\mathrm{C3}|\mathrm{A5B6}) + f(\mathrm{C4}|\mathrm{A5B6}) + f(\mathrm{C4}|\mathrm{A6B5}) + f(\mathrm{C3}|\mathrm{A6B5})} \quad (4)$$

Because we wanted to assess whether changes in attention were resistant to changes in explicit feedback, responses from the first two blocks of trials in the stage were excluded from this analysis.

The mean explicit shift score, $I(\mathrm{A}|\mathrm{X}) - E(\mathrm{A}|\mathrm{Y})$, for all 78 participants was .19. A one-sample Student's t test revealed that this value was significantly different from 0 [$t(77) = 2.84$, $p < .006$, $d = 0.32$]. Thus, the context-specific intradimensional shift was significantly easier than the extradimensional shift, again indicating that people tended to pay attention to the dimension originally relevant in each specific context.

Figure 3 shows the mean magnitude of each of the components of the explicit-shift effect, $I(\mathrm{A}|\mathrm{X})$ and $E(\mathrm{A}|\mathrm{Y})$. To provide some indication of the progression of the effect across the stage, these components are also shown separately for the first two blocks of trials in Stage 3 and for the final two blocks of trials, in which each participant reached the learning criterion for the stage. The overall difference in the magnitudes of both components between the first and last blocks of training [$t(77) = 2.27$, $p = .026$] reflects acquisition of the categorization task, and the absence of a difference between $I(\mathrm{A}|\mathrm{X})$ and $E(\mathrm{A}|\mathrm{Y})$ on the final blocks of training

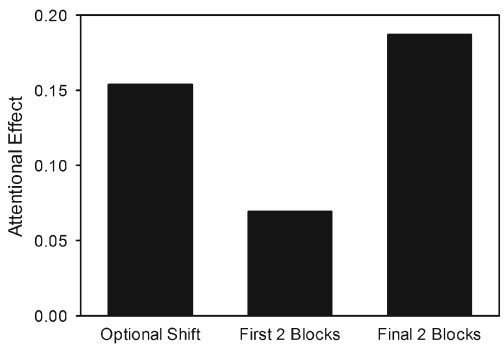

Fig. 2 The mean magnitudes of the optional-shift effect. The first bar shows the magnitude of the effect calculated across Stage 2, with the exception of the first two blocks of trials. The remaining bars show the effect during the first two blocks of trials and during the final two blocks of trials in Stage 2

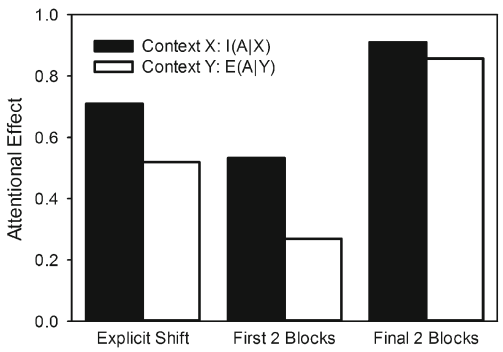

Fig. 3 The mean magnitudes of individual components of the explicit-shift effect in Context X, $I(\mathrm{A}|\mathrm{X})$, and Context Y, $E(\mathrm{A}|\mathrm{Y})$. The overall explicit-shift effect was calculated by taking the difference of these two components. The first two bars show the magnitude of the effect calculated across Stage 3, with the exception of the first two blocks of trials. The remaining bars show the effect during the first two blocks of trials and during the final two blocks of trials in Stage 3

[$t(77) = 1.22$, $p = .22$] is to be expected, since training continued until performance was nearly errorless.

Discussion

This experiment tested whether the amount of attention that is paid to a stimulus may be modulated by the context in which it is presented. Participants first learned to categorize stimuli that were presented in each of two distinct contexts. In Context X, Dimension A signaled category membership while Dimension B was irrelevant, whereas in Context Y, Dimension B signaled category membership while Dimension A was irrelevant.

During the optional shift of Stage 2, participants received no corrective feedback on half of the trials, so that attention to either dimension could yield 100% accuracy. Examination of the responses made on these trials revealed that more attention was paid to Dimension A than to Dimension B in Context X, and more attention was paid to Dimension B than to Dimension A in Context Y.

In the explicit shift of Stage 3, corrective feedback was given on all trials, and Dimension A was relevant to the solution of the categorization task in both Contexts X and Y, thereby constituting an intradimensional shift in Context X and an extradimensional shift in Context Y. The participants made fewer errors when performing the task in Context X than they did on the task in Context Y. These results, together with those from the optional-shift task, are consistent with the proposal that, as a consequence of learning during the Stage 1 categorization task, participants attended more to Dimension A than to Dimension B in Context X, and that they attended more to Dimension B than to Dimension A in Context Y.

These results support previous evidence that the attention paid to a stimulus may be modulated by the context in which that stimulus is presented (e.g., Aha & Goldstone, 1992; Sewell & Lewandowsky, 2011; Sloutsky & Fisher, 2008; Yang & Lewandowsky, 2003). This idea is not consistent with the perspective of conventional models of learning and attention, which largely ignore the role of context. Our results are also not consistent with those of Griffiths and Le Pelley (2009), who found no evidence of context-dependent attention in blocking. In one of their experiments, some cues (denoted "competing" or "blocking") were first established as predictors of a particular outcome [A→O1, C→O1, E→O1]. These cues were subsequently paired with the same outcome in combination with novel "target" or "blocked" cues [AB→O1, CD→O1, EF→O1]. Models of attention in associative learning (e.g., Mackintosh, 1975) predict that, as a consequence of this training, attention should be lower to the target (blocked) cues than to the competing (blocking) cues. In a third stage of training, the target cues were paired with a novel outcome, either in combination with the same competing cue as in the previous stage [AB→O2] or with a different competing cue [ED→O2]. Finally, participants were asked to make a causal rating for each cue. Griffiths and Le Pelley argued that, if attention was modulated by context, participants should attend less to (and therefore learn less about) targets presented in compound with the same competing cue [B] than those presented with different competing cues [D]. Although their results were consistent with participants attending more to the competing cues than to the target cues, they found no effect of context on the ratings given to the target cues.

However, reasons other than an inability of context to modulate attention could explain Griffiths and Le Pelley's (2009) results. In their experiments, each competing cue served as the context for a single target cue, and each target was presented in the presence of a single competing cue (during Stage 2). In our experiment, each of the four stimulus compounds trained in Stage 1 was presented in each of the two contexts. Support for the notion that variation in the context in which a cue is presented might promote context-dependent attention comes from experiments exploring the inverse base-rate effect. Medin and Edelson (1988) trained participants on a task in which cues A and B were paired with one outcome [AB→O1] on a relatively large number of trials, but cues A and C signaled a different outcome [AC→O2] on relatively few trials. When the common cue, A, was presented on test trials, participants predicted outcome O1, consistent with base-rate information. When the other two cues, B and C, were presented in compound, however, outcome O2 was chosen. Kruschke (2009; see also Medin & Edelson, 1988) suggested that these results are best explained in terms of context-dependent attention. He argued that the contingencies between A and O1 and between B and O1 are acquired rapidly, due to the frequency of AB trials. When participants are exposed to AC→O2 trials, however, they selectively learn that C predicts O2. That is, attention to A is dependent on the context in which it is presented: On AB trials, participants attend to both cues, but on AC trials, they attend only to C. In his EXIT model, Kruschke (2001) formalized this idea within a connectionist framework in which the attention paid to each stimulus (or feature) is modulated by a set of exemplar nodes. Hence, in EXIT, it is possible to attend to A when it is presented in compound with B, and not to attend to it when it is presented in compound with C.

Despite its success with the inverse base-rate effect, EXIT (Kruschke, 2001) cannot explain the results of the experiment reported here, because EXIT only attends to individual cues instead of entire dimensions. The model would not, therefore, necessarily predict that there should be any modulation of the attention paid to the novel values

of dimensions. It is, however, possible to conceive of a model based on the same principles as EXIT that could overcome this problem (Kruschke, 2011). Specifically, attentional changes might be applied to an entire stimulus dimension if the stimuli were represented as a pattern of distributed activation over an array of input units. It is reasonable to assume that within such a representational scheme, there would be greater overlap between the patterns of activation evoked by stimuli belonging to the same stimulus dimension than there would be between stimuli belonging to different dimensions. Furthermore, it is likely that representing stimuli in this fashion would allow EXIT to predict contextual modulation of attention using the experimental design shown in Table 1. Such an effect, however, would depend on the similarity of the patterns presented during the different stages of the experiment. Since any attentional modulation within EXIT is exemplar specific, attention will generalize to a novel pattern only to the extent that that pattern activates the same exemplar unit. Hence, in Stage 2 of the experiment, the pattern A3B3 presented in Context X might activate the exemplar node for "A1B1 in Context X" (and, indeed, the exemplar nodes for the other three patterns trained in Context X) sufficiently to modulate attention to Dimensions A and B if these two patterns are sufficiently similar. One might, therefore, predict that by increasing the salience of the contextual cues (and, hence, the similarity of the patterns presented in the same context) one would increase the extent to which the context modulated attention to the stimuli presented within it.

The idea that attention to a stimulus might be modulated by the context in which it is presented is not new. Sutherland and Mackintosh (1971) proposed a two-process model of learning in which stimuli are processed by a number of analyzers, each of which is sensitive to a particular dimension. Responses may be attached to the outputs from these analyzers, and the strength of a response attachment will change as a consequence of reward or nonreward. The size of this change is proportional to the strength of the analyzer. When the outputs of an analyzer make correct predictions about trial outcomes, the strength of that analyzer is increased, and that of all others is decreased. Finally, behavior is controlled by the most active analyzer or analyzers. This model is capable of explaining the ID–ED shift effect. In the first stage of training, the analyzer for the relevant dimension will gain strength, and that for the irrelevant dimension will lose strength. As a consequence, an intradimensional shift will be acquired rapidly, since the appropriate analyzer is strong so that response attachments will be learned rapidly, whereas an extradimensional shift will be acquired less rapidly, since the appropriate analyzer is weak. The model can also explain performance on an optional-shift task, for similar reasons. Sutherland and Mackintosh suggested that the strength of analyzers may be dependent on the context in which a stimulus is presented. They stated that "when an animal learns to switch in a given analyzer, it learns to switch it in in a given situation. The rat that has learned to respond in a jumping stand to black–white differences will not show an increased tendency to control its responses by responding to brightness cues in totally different situations such as its home cage" (p. 55). They did not define precise rules to describe how contextual stimuli might control attention in the type of experimental design that we have described here, but our results are consistent with the principles of Sutherland and Mackintosh's model.

Finally, we must also consider an explanation of our results in terms of models of inductive reasoning. It is possible that our participants simply learned that Dimension A predicted category membership in Context X and that Dimension B predicted category membership in Context Y. Hence, participants' performance on the optional-shift test might have been controlled by a set of learned schema. Kemp, Goodman, and Tenenbaum (2007, 2010) described a Bayesian framework for learning causal schema that could determine the causal type of an object and the interactions between causal types. Within this framework, we might consider a number of different types of cause: items from Dimension A that predict membership of categories C1 and C2 in combination with Context X, or items from Dimension B that predict membership of categories C3 and C4 in combination with Context Y. Such a framework may well be able to account for our data, but we do not believe that computational models of causal induction have been formally applied to category structures of the type that our participants were trained on. This Bayesian rule-induction framework is thematically consistent with notions of selective attention, in that the induced rules selectively encode information from different dimensions in different contexts. Bayesian approaches can also be applied to associative representations, with or without attentional learning (Kruschke, 2006, 2008).

The phenomenon of context-dependent selective attention is not merely one more curiosity in the catalog of findings that theories of learning should address. Instead, the frequent real-world demand for context-dependent selective attention may be the crucial motivation for selective attention in learning at all. In simulations of the evolution of learning mechanisms, it has been shown that environmental structures like Stage 1 of our design, in which different dimensions are relevant in different contexts, are exactly the structures that make rapid attentional learning most adaptive (Kruschke & Hullinger, 2010).

In summary and conclusion, using a variant of an ID–ED shift design, our results add to the evidence that changes in the attention paid to a stimulus dimension over the course of categorization (or discrimination) training can be conditionally modulated by the context. We have also shown, for the

first time, that these changes in attention affect the performance of subsequent categorization tasks. Conventional models of attention in associative learning largely ignore the influence of context and, hence, fail to predict our results. The effect may be explained by a model of learning in which the attention paid to individual cues or dimensions within a compound may be modulated by contextual cues.

Author note D.N.G. is a visiting fellow in the School of Psychology, University of New South Wales, Sydney, Australia. This work was supported by National Science Foundation Grant BCS 9910720, an Experimental Psychology Society Study Visit Grant, and a Royal Society University Research Fellowship. D.N.G. is grateful to the Department of Psychological and Brain Sciences, Indiana University, Bloomington, for the generous provision of facilities while the rationale behind the experiment was being developed.

Appendix: Modeling Sloutsky and Fisher (2008) without attention

We conducted a number of computer simulations to test the predictions of Pearce's (1994) configural theory and Rescorla and Wagner's (1972) elemental model of associative learning concerning Sloutsky and Fisher's (2008) experiment. In all of these simulations, the 16 training and 16 test triads used by Sloutsky and Fisher were represented as unique 14-bit input vectors (see Tables 2 and 3). Each of the three items within a triad (target, Test Item 1, and Test Item 2) was coded over four bits that indicated the presence (1) or absence (0) of each of the two possible shapes (triangle or circle) and colors (black or white) at that location. The final two bits of the vectors represented the presence or absence of each of the two contexts. For the triads for which the right test item (Test Item 1) was correct, the response λ was equal to +1, and for the triads for which the left test item (Test Item 2) was correct, λ was equal to –1. In all simulations, the learning rate parameters, β_E and β_I, were both equal to .1. In each epoch of training, each training triad was presented once in a random order. Associative strengths were updated at the end of each trial. The simulations were allowed to run until the associative strength of each training triad had reached asymptote, after which the net associative strength of each test triad was calculated. When the net associative strength of a test triad was greater than zero, this was judged to be consistent with the selection of Test Item 1 (right response), whereas a net associative strength less than zero was deemed to be consistent with the selection of Test Item 2 (left response).

For simulations of the Rescorla–Wagner (1972) model, unique combinations of features within triads were assumed to generate configurational cues (Wagner & Rescorla, 1972). The net associative strength of an input vector, ΣV, was calculated by taking the sum of the associative strengths of each active feature, context, and configurational cue. At the end of each trial, the associative strength of each of these active elements was updated using Eq. A1:

$$\Delta V_X = \alpha_X \beta(\lambda - \Sigma V), \tag{A1}$$

where V_X is the associative strength of an individual element and α_X is the salience of that element.

Across 19 simulation runs, the saliences of the 12 bits representing the features present at each location and the two bits representing the context were manipulated. In the first 10 simulation runs, α was held at .1 for the features and was varied in the range .01–.1 in increments of .01 for the contexts. For the remaining nine runs, α was held at .1 for the contexts and varied in the range .01–.09 in increments of .01 for the features. In all cases, the model made the correct predictions: The net associative strength of each test triad was consistent with the context-appropriate response. The results of a simulation in which α = .1 for both features and contexts are presented in Table 3 in the column labeled "R-W."

For simulations of configural theory, it was assumed that each bit of the input vector could excite a single input unit in a network. The activation of an individual input unit, u_i, was determined by the intensity of the bit that excited it, I_i, relative to the sum of all bits in the input vector, ΣI, as shown in Eq. A2:

$$u_i = \frac{I_i}{\sqrt{\Sigma I^2}}. \tag{A2}$$

The first time an input vector was presented to the network, a single configural unit, j, was activated maximally, and the strength of the connection between that configural unit and each active input unit, $w_{i,j}$, was set to equal the activation of the input unit, u_i. Whenever an input vector was presented, the activation of each configural unit in the network, a_j, was determined according to Eq. A3, where d is a discriminability parameter (Pearce et al., 2008) that affects the degree of generalization between input vectors.

$$a_j = (\Sigma w_{i,j}.u_i)^d. \tag{A3}$$

The net associative strength of input vector k, V_k, was given by Eq. A4, where E_j is the associative strength of configural unit j:

$$V_k = \Sigma a_j.E_j. \tag{A4}$$

Finally, at the end of each trial, the associative strength of the maximally activated configural unit, E_k, was updated using Eq. A5:

$$\Delta E_k = \beta \cdot (\lambda - V_k). \tag{A5}$$

Configural theory correctly predicted context-appropriate responding whenever the discriminability parameter, d, was

Table 2 The 16 training triads employed by Sloutsky and Fisher (2008) and the 14-bit input vectors used to represent them in computer simulations of two models of associative learning

Triad	Target				Test item 1				Test item 2				Context		Response
	△	○	■	□	△	○	■	□	△	○	■	□	X	Y	
	1	0	0	1	1	0	0	1	0	1	0	1	1	0	+1
	0	1	0	1	0	1	0	1	1	0	0	1	1	0	+1
	1	0	1	0	1	0	1	0	0	1	1	0	1	0	+1
	0	1	1	0	0	1	1	0	1	0	1	0	1	0	+1
	1	0	0	1	0	1	0	1	1	0	0	1	1	0	-1
	0	1	0	1	1	0	0	1	0	1	0	1	1	0	-1
	1	0	1	0	0	1	1	0	1	0	1	0	1	0	-1
	0	1	1	0	1	0	1	0	0	1	1	0	1	0	-1
	1	0	1	0	1	0	1	0	1	0	0	1	0	1	+1
	1	0	0	1	1	0	0	1	1	0	1	0	0	1	+1
	0	1	1	0	0	1	1	0	0	1	0	1	0	1	+1
	0	1	0	1	0	1	0	1	0	1	1	0	0	1	+1
	1	0	1	0	1	0	0	1	1	0	1	0	0	1	-1
	1	0	0	1	1	0	1	0	1	0	0	1	0	1	-1
	0	1	1	0	0	1	0	1	0	1	1	0	0	1	-1
	0	1	0	1	0	1	1	0	0	1	0	1	0	1	-1

greater than 2.0. Over eight sets of 19 simulation runs, the intensities of the features and the contexts, as well as the value of d, were independently manipulated. Within each set of 19 simulations, the intensities of the features and contexts were varied in the same manner as α was in the simulations of the Rescorla-Wagner model, but with values in the range 1–10 in increments of 1. For each set of simulations, d was set to a different value (2.0, 2.1, 2.2, 2.5, 3, 5, 7.5, and 10). When the discriminability parameter was equal to 2.0, the net associative strength of all test triads was close to zero. In all other simulation runs, the net associative strength of the test triads was consistent with Sloutsky and Fisher's results. The final column of Table 3, labeled "CT," shows the results of a simulation of configural theory in which the intensities of all features and contexts were 1 and d was 3.0.

The net associative strength of test triads is a function of generalization from each of the 16 training triads. To understand why Sloutsky and Fisher's (2008) training regimen

Table 3 The 16 test triads employed by Sloutsky and Fisher (2008) and the 14-bit input vectors used to represent them in our computer simulations. The final two columns show the predicted associative strength of each triad following representative simulations of the Rescorla–Wagner (1972) model (R-W) and configural theory (CT)

Triad	*Target*				*Test item 1*				*Test item 2*				*Context*		*ΣV*	
	△	○	■	□	△	○	■	□	△	○	■	□	X	Y	R-W	CT
	1	0	0	1	1	0	1	0	0	1	0	1	1	0	0.33	0.25
	0	1	0	1	0	1	1	0	1	0	0	1	1	0	0.33	0.25
	1	0	1	0	1	0	0	1	0	1	1	0	1	0	0.33	0.25
	0	1	1	0	0	1	0	1	1	0	1	0	1	0	0.33	0.25
	1	0	0	1	0	1	0	1	1	0	1	0	1	0	-0.33	-0.25
	0	1	0	1	1	0	0	1	0	1	1	0	1	0	-0.33	-0.25
	1	0	1	0	0	1	1	0	1	0	0	1	1	0	-0.33	-0.25
	0	1	1	0	1	0	1	0	0	1	0	1	1	0	-0.33	-0.25
	1	0	0	1	1	0	1	0	0	1	0	1	0	1	-0.33	-0.25
	0	1	0	1	0	1	1	0	1	0	0	1	0	1	-0.33	-0.25
	1	0	1	0	1	0	0	1	0	1	1	0	0	1	-0.33	-0.25
	0	1	1	0	0	1	0	1	1	0	1	0	0	1	-0.33	-0.25
	1	0	0	1	0	1	0	1	1	0	1	0	0	1	0.33	0.25
	0	1	0	1	1	0	0	1	0	1	1	0	0	1	0.33	0.25
	1	0	1	0	0	1	1	0	1	0	0	1	0	1	0.33	0.25
	0	1	1	0	1	0	1	0	0	1	0	1	0	1	0.33	0.25

might result in asymmetrical generalization of associative strength to the test triads, let us consider the predictions of configural theory regarding the triad shown in the first row of Table 3. The generalization of associative strength from one input vector to another is determined by the similarity of the two vectors. When $d = 3$ and the intensity of each active bit is 1, Eq. A6 may be used to calculate the similarity of two vectors, A and B, where N_C is the number of features that they have in common, N_A is the number of features in vector A, and N_B is the number of features in vector B.

$$ {}_{A}S_{B} = \left(\frac{N_C}{\sqrt{N_A \times N_B}} \right)^3 \tag{A6} $$

Examination of Table 2 reveals that the training triad most similar to the first test triad has six (out of seven) features in common with it and was associated with the

right response (+1). Other training triads associated with this response share five, four, three, three, three, two, and two features with the first test triad. The eight training triads that were associated with the left response (–1) share five, five, four, four, four, three, two, and one features with that test triad. The average number of features that each set of training triads share with the test triad is, therefore, 3.5. We can see from Eq. A6, however, that the relationship between the proportion of features that two vectors share and their similarity is not linear. Instead, it is described by the function $\cos^3 \theta$, where θ is the inner product of the two vectors. It is this nonlinear relationship that gives rise to the predictions described above. The similarities of two vectors that share six, five, four, three, two, or one feature(s) (out of seven) are .63, .36, .19, .08, .02, and .003, respectively. Hence, for each test triad, the training triad that is most similar to it will have a disproportionate effect on its net associative strength. In each case, this happens to be a triad associated with the context-appropriate response.

References

Aha, D. W., & Goldstone, R. L. (1992). Concept learning and flexible weighting. In J. K. Kruschke (Ed.), *Proceedings of the 14th Annual Conference of the Cognitive Science Society* (pp. 534–539). Hillsdale, NJ: Erlbaum.

Duffaud, A. M., Killcross, 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,* 534–542.

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,* 363–373.

Griffiths, O., & Le Pelley, M. E. (2009). Attentional changes in blocking are not a consequence of lateral inhibition. *Learning & Behavior, 37,* 27–41.

Kemp, C., Goodman, N. D., & Tenenbaum, J. B. (2007). Learning causal schemata. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 389–394). Austin, TX: Cognitive Science Society.

Kemp, C., Goodman, N. D., & Tenenbaum, J. B. (2010). Learning to learn causal models. *Cognitive Science, 34,* 1185–1243.

Kendler, T. S., & Kendler, H. H. (1964). Optional shift behavior of albino rats. *Psychonomic Science, 1,* 5–6.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44. doi:10.1037/0033-295X.99.1.22

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science, 8,* 225–247. doi:10.1080/095400996116893

Kruschke, J. K. (2001). Towards a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45,* 812–863.

Kruschke, J. K. (2006). Locally Bayesian learning with applications to retrospective revaluation and highlighting. *Psychological Review, 113,* 677–699.

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior, 36,* 210–226.

Kruschke, J. K. (2009). Highlighting: A canonical experiment. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 153–185). San Diego, CA: Academic Press.

Kruschke, J. K. (2011). Models of attentional learning. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 120–152). Cambridge, U.K.: Cambridge University Press.

Kruschke, J. K., & Hullinger, R. A. (2010). Evolution of attention in learning. In N. A. Schmajuk (Ed.), *Computational models of conditioning* (pp. 10–52). Cambridge, U.K.: Cambridge University Press.

Mackintosh, N. J. (1974). *The psychology of animal learning*. London: Academic Press.

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

Mackintosh, N. J., & Little, L. (1969). Intradimensional and extradimensional shift learning by pigeons. *Psychonomic Science, 14,* 5–6.

Medin, D. L., & Edelson, S. M. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology. General, 117,* 68–85.

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

Pearce, J. M., Esber, G. R., George, D. N., & Haselgrove, M. (2008). The nature of discrimination learning in pigeons. *Learning & Behavior, 36,* 188–199.

Pearce, J. M., George, D. N., & Redhead, E. S. (1998). The role of attention in the solution of conditional discriminations. In N. A. Schmajuk & P. C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals* (pp. 249–275). Washington, DC: American Psychological Association.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552. doi:10.1037/0033-295X.87.6.532

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). New York, NY: Appleton-Century-Crofts.

Sewell, D. K., & Lewandowsky, S. (2011). Restructuring partitioned knowledge: The role of recoordination in category learning. *Cognitive Psychology, 62,* 81–122.

Sloutsky, V. M., & Fisher, A. V. (2008). Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development, 79,* 639–651.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York, NY: Academic Press.

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301–336). New York, NY: Academic Press.

Yang, L. X., & Lewandowsky, S. (2003). Context-gated knowledge partitioning in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 663–679.