Submitted for publication.
Please do not quote.

Version of April 26, 2001.
Comments welcome.

Cue competition in function learning: Blocking and highlighting

John K. Kruschke
Indiana University, Bloomington

In function learning, people learn to predict a *continuous* outcome from continuous cues. In category learning, people learn to predict a *nominal* outcome. The present research demonstrates that two complementary forms of cue competition, previously found in category learning, also occur in function learning. One form of cue competition is *blocking* of learning about a redundant cue (Kamin, 1968). A second form of cue competition is *highlighting* of a diagnostic cue (a.k.a. the inverse base rate effect; Medin & Edelson, 1988). For tests with conflicting cues, the results show bimodality of responses, as opposed to averaging, which implies exclusive selectivity that cannot be discerned from category learning paradigms. It is argued that these effects are caused, in both category and function learning, by attentional shifts. No previously published model of function learning can account for these effects, but a model by Kalish, Lewandowsky, and Kruschke (2001) is promising.

This article reports evidence of two types of strong cue competition in function learning. One effect is "blocking" of learning about a redundant relevant cue (Kamin, 1968). The other effect is what I call "highlighting," previously referred to as the inverse base rate effect (Medin & Edelson, 1988). There are two main implications of this work. First, because the effects are now evident in both function learning and category learning, the results suggest that future theories of learning in either domain should be designed to address the other domain as well. Second, because theories of these cue competition effects in category learning posit a central role for selective attention (Kruschke, 1996; Kruschke & Blair, 2000), theories of function learning should also incorporate selective attention.

The new data also indicate mutual exclusivity in cue competition, as opposed to averaging, that cannot be measured in category learning paradigms. Consider what might happen when an observer is confronted by two conflicting cues simultaneously. For example, suppose that the person has previously learned that cue A indicates outcome 1, and cue C indicates outcome 3, where outcomes 1 and 3 are values on a metric scale. Suppose that the person is then confronted by a test case in which the conflicting cues A and C are paired. The observer might contemplate the responses for both cues and generate a response that best reflects an average of the two, i.e., response 2. Alternatively, the observer might exclusively select one or the other cue (or cue-outcome link) and generate either response 1 or response 3, perhaps bimodally. Yet another possibility is that the person could just decide that the conflicting cues imply that all bets are off, and so s/he just chooses randomly (uniformly) from the available response options. Still other possibilities exist, of course. My point is that these three different response tendencies, which leave different signatures in the function learning paradigm, cannot be so easily distinguished in the standard category learning paradigm. This is because in category learning there are no intermediate response options between the nominal category labels. The results reported below are most consistent with mutual exclusivity in cue selection, as opposed to averaging.

Background

Function learning is common in everyday tasks. As just a few examples, consider that people can learn how hard to throw a ball in response to visual cues about the distance of the catcher, paramedics can learn how much medication to administer in response to cues about the weight of the patient, and investors can learn how much stock to purchase in response to cues about market trends. Category learning is distinct from function learning only in that the outcomes are categorical values instead of metric values. That is, the outcomes are merely labels (without size or order), instead of magnitudes. This seemingly small difference suggests that similar psychological mechanisms should be involved in both, and that research about the two types of learning should mutually inform each other (see review by Busemeyer, Byun, Delosh, & McDaniel, 1997). One of the central types of phenomena observed in category learning is cue competition, wherein alternative cues apparently compete to gain associative predictiveness of the outcome. It seems reasonable to expect that cue competition should also occur in function learn-

This research was supported in part by NSF grant BCS 9910720. The author thanks Michael Kalish and Stephan Lewandowsky for motivation to pursue this research. For helpful comments on drafts of this article, the author thanks Jerome Busemeyer and Stephan Lewandowsky. For assistance administering the experiment, the author thanks Twanna Allen, Colin Campbell, Dan Hall, Kara Kohnen, Indya Watts and Maxine Weiss.

Correspondence can be addressed to John K. Kruschke, Department of Psychology, 1101 E. 10th St., Indiana University, Bloomington IN 47405-7007, or via electronic mail to kruschke@indiana.edu. The author's world wide web page is at http://www.indiana.edu/~kruschke/

ing. Cue competition is fundamental in associative learning, but only a little work has found cue competition in function learning (e.g. Birnbaum, 1976; Busemeyer, Myung, & McDaniel, 1993; Mellers, 1986).

Perhaps the most well known example of cue competition in category learning is blocking, which has been observed in a variety of species and procedures (e.g. Dickinson, Shanks, & Evenden, 1984; Kamin, 1968; Kruschke & Blair, 2000; Shanks, 1985), and which revolutionized theories of associative learning. In blocking, the participant first learns that a cue, denoted here as A, perfectly predicts an outcome. In subsequent training, the participant experiences cases of cue A paired with another cue, denoted B, still perfectly predictive of the outcome. Despite the fact that B perfectly predicts the outcome, people apparently do not learn to strongly associate it with the outcome. That is, learning about B has been blocked by previous learning about A. To my knowledge, blocking has never been sought in the context of function learning, yet it should exist if blocking is truly a fundamental and pervasive phenomenon in learning.

Another dramatic case of cue competition in category learning is what I call highlighting, previously referred to as the inverse base rate effect (Medin & Edelson, 1988). In this paradigm, people initially learn that a pair of cues, denoted I.E, indicate outcome O_E. In the next phase of training, new cases are introduced in which cue pair I.L indicates outcome O_L. Thus, cue I is an imperfect predictor, while cue E is a perfect predictor of outcome O_E and cue L is a perfect predictor of outcome O_L. Despite the symmetry of each outcome having one shared cue and one perfectly predictive cue, people seem to learn a stronger association from L to O_L than from E to O_E, and a stronger association from I to O_E than from I to O_L. This is evidenced by the fact that people tend to respond with O_E when tested on I alone, and tend to respond with O_L when tested on the pair E.L. This effect, wherein cue L appears to be selectively highlighted when learning I.L$\rightarrow O_L$ (and cue I is selectively ignored when learning I.L$\rightarrow O_L$) has been documented in a variety of circumstances (Dennis & Kruschke, 1998; Fagot, Kruschke, Depy, & Vauclair, 1998; Kalish & Kruschke, 2000; Kruschke, 1996, 2002; Medin & Bettger, 1991; Shanks, 1992). Highlighting is complementary to blocking, because the second cue (L) experiences enhanced learning rather than diminished learning. Like blocking, highlighting has never been sought in the context of function learning, yet it should exist if it reflects a fundamental cue competition effect.

Experiment: blocking and highlighting

Table 1 shows the design of the blocking component of the experiment. The leftmost column indicates the phase of the blocking component; there were two training phases followed by a test phase. Each row of the training phases shows a single training trial. For example, the row with cue A3 and outcome 3 indicates that the third level of cue A occurred with outcome level 3.

The correct outcome level depended on the group to which the participant was assigned. Half the participants learned positive functions of the cues, and the other half learned negative functions. These outcomes are indicated in the two rightmost columns of Table 1. This manipulation is important because, if analogous effects are observed for both positive and negative functions, then we have evidence that the effects are due to *learned* functional relationships, and are not an artifact of previously existing biases toward positive linear relationships (e.g. Sawyer, 1991; Summers, Summers, & Karkau, 1969).

Table 1
Design of the blocking component of the experiment.

		Correct Response	
Phase	Cue(s)	Positive Function Group	Negative Function Group
Train 1	A3	3	6
	A4	4	5
	A5	5	4
	A6	6	3
Train 2	A3.B3	3	6
	A4.B4	4	5
	A5.B5	5	4
	A6.B6	6	3
	Ca3.Cb3	3	6
	Ca4.Cb4	4	5
	Ca5.Cb5	5	4
	Ca6.Cb6	6	3
Test	2nd phase training items, plus:		
	B1.Ca8, B1.Cb8,	A1.Ca8, A1.Cb8	
	B3.Ca6, B3.Cb6,	A3.Ca6, A3.Cb6	
	B6.Ca3, B6.Cb3,	A6.Ca3, A6.Cb3	
	B8.Ca1, B8.Cb1,	A8.Ca1, A8.Cb1	

Note. A1-A8 denote the eight levels of the cue that acted as the blocker. B1-B8 denote the eight levels of the cue that was blocked. Ca1-Ca8 and Cb1-Cb8 denote the eight levels of the two control cues. A dot (period) separates two cues that appeared together.

In the blocking design (Table 1), notice that cue B is the blocked cue, completely redundant with cue A but also perfectly predictive of the outcome. Cues Ca and Cb are control cues, which occur as often as cue B. Cues Ca and Cb are structurally equivalent in the design, and so the results will be collapsed across occurrences of Ca and Cb and denoted simply as cue C. The various items were presented in random order across trials of training.

If blocking occurs, then responses to cue B will be weak compared to the control cues C. This is assessed in the test phase by presenting people with conflicting cues, such as B3.C6. (B3.C6 denotes cases of B3.Ca6 and B3.Cb6.) If cue B has been blocked, then the control cue will dominate, and the responses should tend to be level 6 and not level 3.

Table 2 shows the design for the highlighting component of the experiment. All participants did both the blocking

Table 2
Design of the highlighting component of the experiment.

		Correct Response	
Phase	Cue(s)	Positive Function Group	Negative Function Group
Train 1	I3.E3	3	6
	I4.E4	4	5
	I5.E5	5	4
	I6.E6	6	3
Train 2	I3.E3	3	6
	I4.E4	4	5
	I5.E5	5	4
	I6.E6	6	3
	I3.L3	6	3
	I4.L4	5	4
	I5.L5	4	5
	I6.L6	3	6
Test	2nd phase training items, plus: E1, L1, I1, E1.L1; E3, L3, I3, E3.L3; E6, L6, I6, E6.L6; E8, L8, I8, E8.L8		

Note. I1-I8 denote the eight levels of the imperfectly predictive cue. E1-E8 denote the eight levels of the earlier learned perfectly predictive cue. L1-L8 denote the eight levels of the later learned perfectly predictive cue. A dot (period) separates two cues that appeared together.

and highlighting components, in counterbalanced order. Like the blocking component, in the highlighting component there were two phases of training followed by a testing phase. Also as in the blocking component, half the participants learned positive functions, and the other half learned negative functions.

In Table 2, notice in training phase 2 that cue I is imperfectly predictive of the outcome. For example, in the positive function group, when I3 occurs with E3, it indicates outcome 3, but when I3 occurs with L3, it indicates outcome 6. This is true for every level of cue I, such that half the time it indicates one outcome level, and half the time the complementary outcome level. Both cues E and L, on the other hand, are perfectly predictive. They indicate opposite sign functions, however. Thus, for participants in the positive function group, cue E3 indicates outcome 3 but cue L3 indicates outcome 6.

In both the blocking and highlighting components, people were trained on the middle four levels of an eight-level scale. The test phases probed levels above and below the training levels, thereby checking for extrapolation of the linear function relating the cues and the responses. The extrapolation tests are especially important, because they test that people actually learned a functional relationship between cue and outcome, as opposed to merely memorizing the particular cue-outcome correspondences experienced in training. If tests of the values outside the training range do not show extrapolation, then any observation of blocking or highlighting could be explained away as just the same old blocking and highlighting already known to occur in category learning. The importance of extrapolation in function learning has been pointed out by other authors in other contexts, both in function learning and in category learning (e.g. Delosh, Busemeyer, & McDaniel, 1997; Erickson & Kruschke, 1998; Kalish et al., 2001).

Method

Participants

Fifty seven students from introductory psychology courses at Indiana University participated for partial course credit. Ages ranged from 17 to 52 years, with a median of 19. There were 33 females and 24 males.

Apparatus and stimuli

Participants were trained individually in dimly lit, sound-dampened cubicles. They sat in front of a standard desktop computer at a comfortable viewing distance, and made responses by pressing keys on the standard keyboard.

The cues were verbal-numerical, and displayed as the cue name and value; e.g., “Body Temp: 95.0 deg.F”. The seven cues included the following: Body Temp, from 95.0 deg.F to 102.0 in increments of 1.0; GPA, from 0.5 to 4.0 in increments of 0.5; Mile Time, from 5:30 to 9:00 in increments of 0:30; Income, from $28,000 to 70,000 in increments of 6,000; Sound Threshold, from 2.0 dB to 23.0 in increments of 3.0; Hair Length, from 2.0 cm to 37.0 in increments of 5.0; Shoe Size, from 5+1/2 to 12+1/2 in increments of 1. Therefore, there were 8 levels of each cue. The response was a “mystery meter” level from 1 to 8, made by pressing the corresponding key on the top row of the keyboard.

Procedure

The designs of the blocking and highlighting components are shown in Tables 1 and 2. The seven cue terms were randomly assigned to abstract cue types, independently for every participant. The ordering of blocking and highlighting was counterbalanced across subjects, as was the assignment to negative or positive functions for each component. Thus, there were eight possible orderings (counterbalanced across subjects): blocking positive function followed by highlighting positive function; highlighting positive followed by blocking positive; blocking positive function followed by highlighting negative function; highlighting negative followed by blocking positive; etc. In every phase of the experiment, the various items were randomly permuted within each block. In the highlighting procedure, there were 10 repetitions of Train 1 and 15 repetitions of Train 2. The Test phase included one trial of each of the training items[1],

[1] For some participants, the highlighting *test* phase inadvertently included more I.E items than I.L items (because of a vestige from early experiments that used different base rates for I.E and I.L). This

one trial of each of the E and L items, and two trials each of the I and E.L items. In the blocking procedure, there were 10 repetitions of Train 1, 10 repetitions of Train 2, and in the test phase there was one trial of each of the training items plus one trial of each of the test items. This total of 336 trials proceeded seamlessly without interruption after the initial instructions. The experiment took about 30 minutes.

The initial instructions, written on the computer screen, stated that participants would see information about a hypothetical person, and would need to predict the level of a "mystery meter" that measured something about the person. The instructions told the participant, "You simply need to observe the information about each person and learn to predict the mystery meter level." On a trial, the cues would appear, one per line, centered horizontally on the screen. When there were two cues, their location (top or bottom) was random on each trial. After a prediction was made, the cues stayed on the screen along with feedback that indicated the correct answer. Feedback to erroneous responses was accompanied by a brief buzzing sound. On test trials without correct answers, the feedback indicated that no official meter level was yet available but their prediction was recorded. The participant pressed the space bar to move to the next trial.

Results

Blocking, positive function group

Because I am interested in effects of learning, data analysis excluded participants who did not learn to respond at accuracy significantly above chance. For the blocking design, I wanted to be sure that people had learned the control cues, which occurred later in training than the blocker. I set chance at 25% correct because only 4 outcomes had occurred during the training phases. To reject the null hypothesis of chance performance on the control-cue items in the last two blocks of training, a person must get at least 5 of 8 occurrences correct.[2] Only 1 participant of 29 in this condition failed to achieve this level of accuracy, leaving $N = 28$.

In the testing phase, performance continued to be highly accurate on the training items, averaging 91% for the blocking cues and 85% for the control cues. Complete results are shown in the Appendix in Table 3.

Figure 1 shows results for novel cue combinations in the test phase. Each row in the figure corresponds to a particular cue combination. The width of the circles in each row indicates the percentage of responses at each of the eight outcome levels. The lower panel of Figure 1 shows that there was robust blocking, with the response appropriate for the control cue dominating the response appropriate to the blocked cue. For example, with test cues B6.C3, the circle at response level 3 is much larger than the circle at response level 6. There was distinct extrapolation of control cue dominance beyond the training range. For example, with test cues B8.C1, the modal response is level 1, which was never experienced during training.

Inferential statistics bolster these descriptive observations that blocking occurred and was extrapolated. Blocking occurred at the scale levels of the training items, insofar as the frequency of "5" or "6" responses to B3.C6 and "3" or "4" responses to B6.C3 was significantly greater than the frequency of "3" or "4" responses to B3.C6 and "5" or "6" responses to B6.C3, $\chi^2_{(df=1,N=104)} = 6.50$, $p < .012$. There was also extrapolation of blocking, in that the frequency of responding with extreme (never trained) scale values ("1" or "2" responses to B8.C1 and "7" or "8" responses to B1.C8) was significantly higher than would be expected by chance responding across the entire scale, $\chi^2_{(df=1,N=112)} = 13.76$, $p < .001$.

The top panel of Figure 1 shows that despite the control cues' dominance of the blocked cue B, the control cues did not dominate the blocking cue A. (In the terminology of asso-

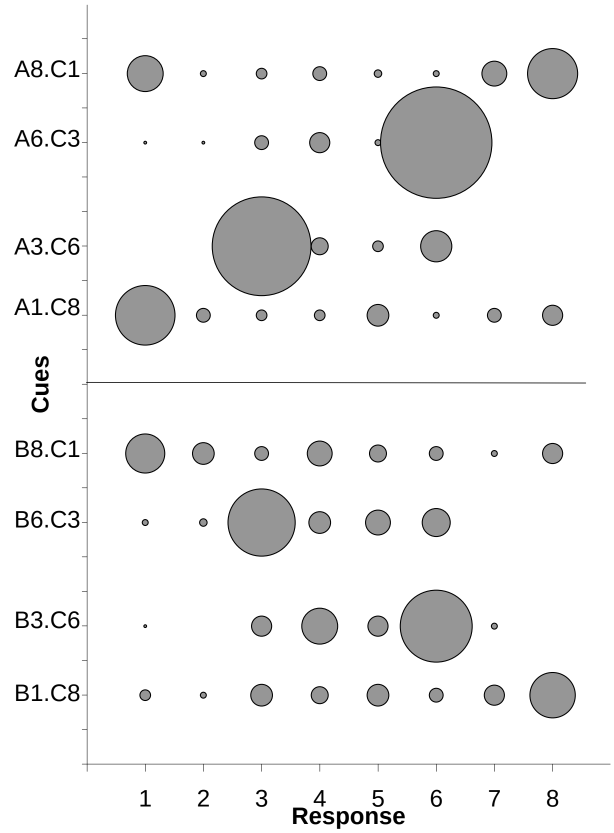

Figure 1. Response percentages for A.C and B.C test trials of blocking, positive function group. Diameter of circle corresponds to response percentage. See the Appendix for numerical values.

procedural infelicity was detected and corrected for the remaining participants. The increased exposure to I.E items during test only makes the results more remarkable, because the preference for the later-trained responses on the conflicting cues E.L must overcome this extra training with I.E.

[2] An alternative measure of learning accuracy is mean absolute deviation between correct and selected response levels. The "hit/miss" measure was used instead because of its simplicity and conservativeness.

ciative learning theory, this can be described as the two control cues mutually *overshadowing* each other during learning, although the effect might also be due in part to the greater overall exposure to A.) For example, with cue combination A6.C3, the modal response is level 6. The modal response in each case follows the linear function appropriate to the blocking cue A, even in the extrapolation region.

Inferential statistics concur with this description of the blocking cue's dominance and extrapolation. At the scale levels of the training items, the frequency of "5" or "6" responses to A6.C3 and "3" or "4" responses to A3.C6 was significantly greater than the frequency of "3" or "4" responses to A6.C3 and "5" or "6" responses to B3.C6, $\chi^2_{(df=1,N=110)} = 14.24$, $p < .001$. There was also extrapolation of blocker dominance, in that the frequency of responding with extreme (never trained) scale values ("1" or "2" responses to A1.C8 and "7" or "8" responses to A8.C1) was significantly higher than would be expected by chance responding across the entire scale, $\chi^2_{(df=1,N=112)} = 29.76$, $p < .001$.

Blocking, negative function group

The results for the negative-function version of blocking were entirely analogous to the positive-function version. The control cues dominated the blocked cue, and the blocking cue dominated the control cues. There was extrapolation of the dominance into the untrained regions of the scales. Thus, the results from the positive-function version cannot be attributed to prior biases about positive functions. Complete details are presented in the Appendix.

Highlighting, positive function group

To exclude non-learners, the same criterion was used as for the blocking components: Participants had to show significantly above chance performance on the later-trained items in the last two blocks of training. This means at least 5 of 8 correct on the I.L trials in the last two blocks of training. Of 29 participants, 4 failed to reach this criterion, yielding N=25 for subsequent analyses.

Performance on the training items continued to be good in the test phase, averaging 98% correct on the I.E items and 88% correct on the I.L items.

Figure 2 shows results from some of the critical test cases. (Table 5 in the Appendix shows complete results.) Responses to cue I by itself showed a clear preference for the early-learned outcome. For example, in the upper panel of Figure 2, it can be seen that the modal response to cue I6 was outcome level 6 (which was the early-learned outcome). The linear trend continues into the extrapolation regions.

Importantly, responses to conflicting cues E.L showed a preference for the later learned outcome. For example, in the lower panel of Figure 2, it can be seen that the modal response to cue combination E6.L6 was outcome level 3, which corresponds with cue L6. There is clear extrapolation of cue L's dominance into untrained extremities of the scales. One aspect worth emphasizing is that the dominant cue, L, indicated a *negative* function, which is a functional

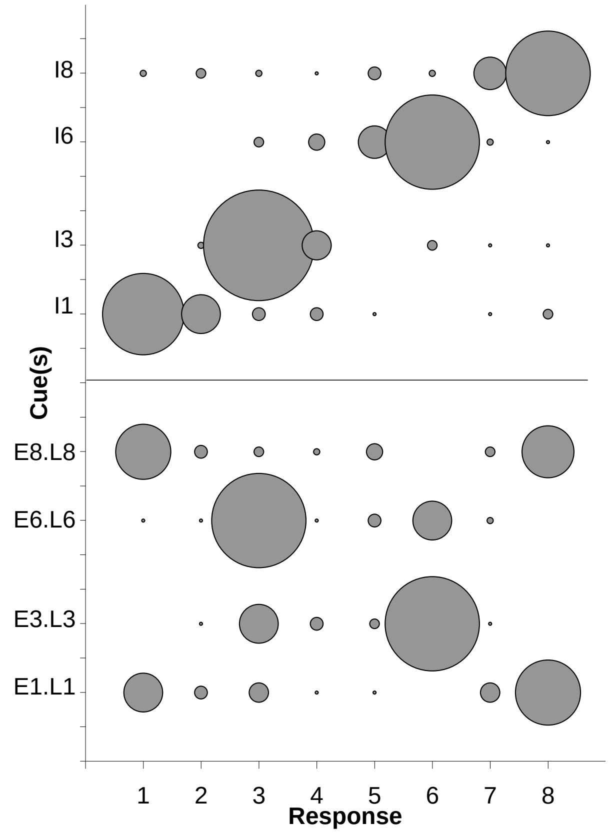

Figure 2. Response percentages for I and E.L test trials of highlighting, positive function group. Diameter of circle corresponds to response percentage. See the Appendix for numerical values.

form known to be difficult to learn and apply (e.g. Brehmer, 1971, 1973). Its dominance over the positive cue is therefore quite remarkable.

Interestingly, responses to the conflicting cues E.L showed a weaker, secondary peak along the function for the early learned outcome. In the lower panel of Figure 2 this can be seen as a negative linear function connecting the largest circles in each row, but a positive linear function connecting the second-largest circles in each row. There are relatively few responses off of these two linear functions. Thus, for the conflicting cues, it seems that participants tended to select cue L to dictate responding, but occasionally selected cue E to dictate responding. It seems that people rarely if ever selected both cues and averaged their predicted outcomes. This observation could not have been found in traditional category learning experiments, because there are no intermediate response levels between categories.

Inferential statistics support these observations. For cue I by itself, the frequency of "1" or "2" responses to I1 and "7" or "8" responses to I8 was greater than uniform responding across the scale range, $\chi^2_{(df=1,N=100)} = 122.9$, $p < .001$. The frequency of "3" or "4" responses to I3 and "5" or "6" responses to I6 was greater than the frequency of "5" or "6" responses to I3 and ”3” or ”4” responses to I6, $\chi^2_{(df=1,N=93)} = 54.2$, $p < .001$. For conflicting cues E.L, the frequency of responding along the negative diagonal was significantly greater than the frequency of responding along the positive diagonal[3], $\chi^2_{(df=1,N=176)} = 10.0$, $p < .001$.

Highlighting, negative function group

The results for the negative-function version of the highlighting component were entirely analogous to the results from the positive-function version. For cue I by itself, the early-learned (negative function) outcome was dominant. For cue combination E.L, the later-learned (positive function) outcome was dominant. In both cases there was clear extrapolation to the untrained regions of the scales. Thus, the results from the positive function version cannot be attributed to prior biases about positive functions, or to some prior “distinctiveness” of negative functions. Complete data are presented in the Appendix.

Conclusion

The results showed strong evidence of blocking and highlighting in function learning. This is the first time that these effects of cue competition from category learning have been found in function learning. Previous researchers have demonstrated forms of cue competition in function learning (e.g. Birnbaum, 1976; Busemeyer et al., 1993; Mellers, 1986), but they have not shown bridges from these particular dramatic cue competition effects in category learning.

The results also showed definite extrapolation to novel values above and below the trained domain. Hence the blocking and highlighting reported here cannot be explained as merely a conflagration of several instances of category learning. It appears, instead, that people really did learn functional relationships between cues and outcomes, and that there was strong competition between the cues.

There was also bimodality in responses for some test cases that used conflicting cues. In particular, for cases of E.L in the highlighting design, the main trend was for responses to follow the function appropriate to cue L, but there was also a weaker peak along the function appropriate to the cue E. That is, people appeared to select primarily cue L to dominate the response, or secondarily cue E. There was not a strong tendency to select both cues and average their implied responses. This bimodality indicates exclusive selection of cues that cannot be discerned from previous demonstrations of highlighting in category learning paradigms, because category learning paradigms do not have response options that are intermediate between the trained outcomes.

Theories of blocking and highlighting in category learning posit a central role for attentional shifting and learning (Kruschke, 1996; Kruschke & Blair, 2000; Kruschke, 2001). The present results add to evidence that selective attention plays a significant role in function learning. The importance of cue selection has been recognized by many researchers of function learning (e.g. Klayman, 1988), but the new data presented in this article make the connection to attentional selection in category learning that much stronger.

There have been a number of thoughtful and successful models of function learning previously described in the literature, but none of them, to my knowledge, can account for the present results. The adaptive regression model (ARM) of Koh (1993), for example, has many attractive properties, and because of its error-driven learning algorithm might be anticipated to be able to show blocking in function learning, analogously to how the error-driven Rescorla-Wagner model (1972) can show blocking in category learning. There is space in this article only for a cursory indication of ARM’s predictions, derived here by armchair analysis, not by numerical simulation. Whereas ARM can, appropriately, learn a larger regression coefficient for the blocking cue than for the blocked cue, the model does not learn about the control cues appropriately. ARM learns half-strength coefficients for each of the two control cues (assuming that the two control cues are equally salient). This causes ARM to generate half-strength outputs for the conflicting-cue test cases, unlike the full-strength modal responses produced by people. More dramatically, ARM is unable to produce the highlighting effect, even qualitatively, because it predicts that the unpredictive cue I should converge to a regression coefficient of zero, and the two predictive cues, E and L, should converge to equal strength (but opposite sign) regression coefficients.

The extrapolation association model (EXAM; Delosh et al., 1997) has been shown to fit a wide variety of phenomena in function learning (for a review, see Busemeyer et al., 1997). EXAM can be thought of as an error-driven exemplar based model (ALCOVE; Kruschke, 1992) with an additional response mechanism that extrapolates linearly from learned exemplars. Armchair analysis suggests that EXAM *can* show, at least qualitatively, the results from the blocking component of the experiment. But EXAM is not able to exhibit highlighting. The exemplar representation and error-driven learning algorithm makes the model respond ambivalently to ambiguous cues. This failure to capture highlighting in function learning is a direct extension of ALCOVE’s (Kruschke, 1992) failure to capture highlighting in category learning.

Kruschke (e.g., 1996, 2002) has argued that highlighting is best explained by rapid shifts of attention during learning. In the early phase of learning, the shared cue I is associated with the early functional outcome. In the later phase of learning, when studying cue combinations I.L, attention shifts to the distinctive cue L, in order to reduce interference from the previously learned knowledge about cue I. In this way, cue I remains associated with the earlier learned outcome, while cue L becomes strongly associated with the later learned outcome. Kruschke and Blair (2000; see also Kruschke, 2001)

[3] The frequency of responding along the positive diagonal was computed as the frequency of responding ”1” or ”2” to E1L1, ”3” or ”4” to E3L3, ”5” or ”6” to E6L6, and ”7” or ”8” to E8L8. The frequency of responding along the negative diagonal was computed as the frequency of responding ”1” or ”2” to E8L8, ”3” or ”4” to E6L6, ”5” or ”6” to E3L3, and ”7” or ”8” to E1L1.

argued that blocking also involves shifts of attention during learning. In the second phase of training in the blocking procedure, people learn to shift attention away from the blocked cue B, toward the blocking cue A, because doing so takes advantage of the already learned association from cue A.

The only extent model of function learning that has this sort of attention shifting ability is a variant of the mixture-of-experts approach used in category learning (Kruschke & Erickson, 1994; Erickson & Kruschke, 1998; Kruschke, 2001), being developed by Kalish et al. (2001). In this approach, there is a population of (linear) functions available for mapping cues to outcomes. These functions can be selectively attended to, in a manner dictated by rapid error reduction. Thus, when learning about cues I.L, the model reduces error by selectively attending to cue L. This shift of attention is learned, so that in subsequent tests with cues E.L, there is greater attention paid to cue L. Future research will need to explore this potential in detail.

References

Birnbaum, M. H. (1976). Intuitive numerical prediction. *American Journal of Psychology*, *89*, 417–429.

Brehmer, B. (1971). Subjects' ability to use functional rules. *Psychonomic Science*, *24*, 259–260.

Brehmer, B. (1973). Single-cue probability learning as a function of the sign and magnitude of the correlation between cue and criterion. *Organizational Behavior and Human Performance*, *9*, 377–395.

Busemeyer, J. R., Byun, E., Delosh, E. L., & McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 405–437). Cambridge, MA: MIT Press.

Busemeyer, J. R., Myung, I. J., & McDaniel, M. A. (1993). Cue competition effects: Empirical tests of adaptive network learning models. *Psychological Science*, *4*, 190–195.

Delosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 968–986.

Dennis, S., & Kruschke, J. K. (1998). Shifting attention in cued recall. *Australian Journal of Psychology*, *50*, 131–138.

Dickinson, A., Shanks, D. R., & Evenden, J. L. (1984). Judgement of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *36A*, 29–50.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107–140.

Fagot, J., Kruschke, J. K., Depy, D., & Vauclair, J. (1998). Associative learning in baboons (papio papio) and humans (homo sapiens): species differences in learned attention to visual features. *Animal Cognition*, *1*, 123–133.

Kalish, M. L., & Kruschke, J. K. (2000). The role of attention shifts in the categorization of continuous dimensioned stimuli. *Psychological Research*, *64*, 105–116.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2001). Population of linear experts:. (In preparation)

Kamin, L. J. (1968). 'Attention-like' processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9–33). Coral Gables, FL: University of Miami Press.

Klayman, J. (1988). On the how and why (not) of learning from outcomes. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 115–162). Amsterdam: North-Holland.

Koh, K. (1993). Induction of combination rules in two-dimensional function learning. *Memory & Cognition*, *21*, 573-590.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22–44.

Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *22*, 3–26.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, ****, **–**. (in press)

Kruschke, J. K. (2002). Eliminative inference does not explain the inverse base rate effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, ****, **–**. (in press)

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review*, *7*, 636–645.

Kruschke, J. K., & Erickson, M. A. (1994). Learning of rules that have high-frequency exceptions: New empirical data and a hybrid connectionist model. In *The proceedings of the sixteenth annual conference of the cognitive science society* (pp. 514–519). Hillsdale, NJ: Erlbaum.

Medin, D. L., & Bettger, J. G. (1991). Sensitivity to changes in base-rate information. *American Journal of Psychology*, *104*, 311–332.

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.

Mellers, B. A. (1986). Test of a distributional theory of intuitive numerical prediction. *Organizational Behavior and Human Decision Processes*, *38*, 279–294.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning: Ii. current research and theory* (p. 64-99). New York: Appleton-Century-Crofts.

Sawyer, J. E. (1991). Effects of risk and ambiguity on judgments of contingency relations and behavioral resource allocation decisions. *Organizational Behavior and Human Performance*, *49*, 124–150.

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology*, *37B*, 1–21.

Shanks, D. R. (1992). Connectionist accounts of the inverse base-rate effect in categorization. *Connection Science*, *4*, 3–18.

Summers, S. A., Summers, R. A., & Karkau, V. T. (1969). Judgments based on different functional relations between interacting cues and criterion. *American Journal of Psychology*, *82*, 203-211.

Table 3
Choice percentages from test phase of blocking component, positive function group.

	Response							
Cues	1	2	3	4	5	6	7	8
A3.B3	0	0	**86**	7	7	0	0	0
A4.B4	0	0	4	**89**	4	4	0	0
A5.B5	0	0	4	0	**89**	7	0	0
A6.B6	0	0	0	0	0	**100**	0	0
Ca3.Cb3	0	4	**82**	4	11	0	0	0
Ca4.Cb4	0	4	11	**75**	7	4	0	0
Ca5.Cb5	0	0	4	7	**89**	0	0	0
Ca6.Cb6	0	0	0	4	4	**93**	0	0
A1.C8	**38**	9	7	7	14	4	9	13
A3.C6	0	0	**63**	11	7	20	0	0
A6.C3	2	2	9	13	4	**71**	0	0
A8.C1	23	4	7	9	5	4	16	**32**
B1.C8	7	4	14	11	14	9	13	**29**
B3.C6	2	0	13	23	13	**46**	4	0
B6.C3	4	5	**43**	14	16	18	0	0
B8.C1	**25**	14	9	16	11	9	4	13

Note. A1-A8 denote the eight levels of the cue that acted as the blocker. B1-B8 denote the eight levels of the cue that was blocked. Ca1-Ca8 and Cb1-Cb8 denote the eight levels of the two control cues. A dot (period) separates two cues that appeared together. Numbers in **bold font** indicate the modal response in each row.

Table 4
Choice percentages from test phase of blocking component, negative function group.

	Response							
Cues	1	2	3	4	5	6	7	8
A3.B3	0	0	6	0	6	**88**	0	0
A4.B4	0	0	0	6	**94**	0	0	0
A5.B5	0	0	0	**88**	12	0	0	0
A6.B6	0	0	**94**	0	6	0	0	0
Ca3.Cb3	0	0	6	0	6	**88**	0	0
Ca4.Cb4	0	0	0	12	**76**	6	6	0
Ca5.Cb5	0	6	6	**82**	6	0	0	0
Ca6.Cb6	0	0	**94**	0	6	0	0	0
A1.C8	24	12	6	12	0	0	15	**33**
A3.C6	0	0	26	12	6	**53**	0	3
A6.C3	0	3	**62**	3	12	21	0	0
A8.C1	**47**	6	3	9	3	9	0	24
B1.C8	**35**	3	15	15	0	9	12	12
B3.C6	3	0	**38**	24	6	18	6	6
B6.C3	0	0	21	18	18	**44**	0	0
B8.C1	15	9	12	6	9	3	12	**35**

Note. A1-A8 denote the eight levels of the cue that acted as the blocker. B1-B8 denote the eight levels of the cue that was blocked. Ca1-Ca8 and Cb1-Cb8 denote the eight levels of the two control cues. A dot (period) separates two cues that appeared together. Numbers in **bold font** indicate the modal response in each row.

Appendix: Complete Data

Blocking, positive function group

Table 3 shows the results from the blocking component for participants who were trained with a positive function.

Blocking, negative function group

It is well known that negative functions are more difficult to learn than positive functions (e.g. Brehmer, 1971, 1973). Therefore it is not surprising that fewer people passed the learning criterion for the negative function version of the blocking design than the positive function version. Of 28 participants, 11 failed to reach the learning criterion (5 correct out of 8 control trials in the final two training blocks). Thus, $N = 17$ in the data reported.

Table 4 shows the results from the test phase. Performance continued to be highly accurate on the training items, averaging 91% for the blocking cues (denoted A3B3-A6B6) and 85% for the control cues (denoted Ca3Cb3-Ca6Cb6).

For the tests of blocking, involving a blocked cue (denoted B) and a control cue (denoted C), two effects stand out, directly analogous to the positive function results. First, there was robust blocking, with the response appropriate for the control cue dominating the response appropriate to the blocked cue. Second, there was distinct extrapolation of the control cue beyond the training range. These two effects can be seen clearly in Table 4, where the modal response for each cue combination, set in bold font, follows exactly the (negative) linear function appropriate to the control cue.

Inferential statistics bolster these descriptive observations that blocking occurred and was extrapolated. Blocking occurred at the scale levels of the training items, insofar as the frequency of ”3” or ”4” responses to B3C6 and ”5” or ”6” responses to B6C3 was significantly greater (because of the negative function learned) than the frequency of ”5” or ”6” responses to B3C6 and ”3” or ”4” responses to B6C3, $\chi^2_{(df=1,N=63)} = 7.0$, $p < .01$. There was also extrapolation of blocking, in that the frequency of responding with extreme (never trained) scale values (”1” or ”2” responses to B1C8 and ”7” or ”8” responses to B8C1) was significantly higher than would be expected by chance responding across the entire scale, $\chi^2_{(df=1,N=68)} = 11.29$, $p < .001$.

Also evident in Table 4 is the dominance of the blocker (denoted A) over the control cues. The modal response in each case, set in bold font in Table 4 follows the (negative) linear function appropriate to the blocker.

Inferential statistics concur with this description of the blocker’s dominance and extrapolation. For scale levels in the training range, the frequency of ”3” or ”4” responses to A6C3 and ”5” or ”6” responses to A3C6 was significantly greater than the frequency of ”3” or ”4” responses to A3C6 and ”5” or ”6” responses to A6C3, $\chi^2_{(df=1,N=66)} = 4.90$,

Table 5
Choice percentages from test phase of highlighting component, positive function group.

	Response							
Cue(s)	1	2	3	4	5	6	7	8
I3.E3	2	0	**96**	2	0	0	0	0
I4.E4	0	0	2	**96**	2	0	0	0
I5.E5	0	0	0	0	**100**	0	0	0
I6.E6	0	0	0	0	0	**98**	0	2
I3.L3	0	4	8	0	0	**88**	0	0
I4.L4	0	0	0	20	**80**	0	0	0
I5.L5	0	0	0	**92**	4	4	0	0
I6.L6	0	0	**92**	0	4	4	0	0
E1	**48**	28	0	8	8	0	0	8
E3	0	0	**76**	12	4	8	0	0
E6	4	0	0	8	0	**72**	8	8
E8	4	0	4	0	8	8	16	**60**
L1	12	4	0	0	8	0	12	**64**
L3	0	4	0	8	8	**76**	4	0
L6	0	0	**80**	4	4	4	4	4
L8	**52**	32	0	0	0	0	4	12
I1	**50**	24	8	8	2	0	2	6
I3	0	4	**68**	18	0	6	2	2
I6	0	0	6	10	20	**58**	4	2
I8	4	6	4	2	8	4	20	**52**
E1.L1	24	8	12	2	2	0	12	**40**
E3.L3	0	2	24	8	6	**58**	2	0
E6.L6	2	2	**58**	2	8	24	4	0
E8.L8	**34**	8	6	4	10	0	6	32

Note. I1-I8 denote the eight levels of the imperfectly predictive cue. E1-E8 denote the eight levels of the earlier learned perfectly predictive cue. L1-L8 denote the eight levels of the later learned perfectly predictive cue. A dot (period) separates two cues that appeared together. Numbers in **bold font** indicate the modal response in each row.

Table 6
Choice percentages from test phase of highlighting component, negative function group.

	Response							
Cue(s)	1	2	3	4	5	6	7	8
I3.E3	0	0	11	2	7	**80**	0	0
I4.E4	0	0	2	7	**89**	2	0	0
I5.E5	0	0	4	**84**	9	2	0	0
I6.E6	0	0	**87**	4	2	4	2	0
I3.L3	0	11	**63**	5	5	16	0	0
I4.L4	0	0	5	**84**	11	0	0	0
I5.L5	0	0	5	5	**89**	0	0	0
I6.L6	0	0	11	16	0	**74**	0	0
E1	5	16	0	11	11	11	16	**32**
E3	0	11	5	26	11	**47**	0	0
E6	0	5	**53**	11	11	16	0	5
E8	*21*	21	11	0	21	5	16	5
L1	**47**	21	0	5	0	16	0	11
L3	0	11	**68**	5	11	0	0	5
L6	0	0	5	16	11	**63**	5	0
L8	0	5	5	5	11	16	5	**53**
I1	5	18	8	3	8	11	18	**29**
I3	0	5	16	13	13	**53**	0	0
I6	0	5	**47**	24	8	8	3	5
I8	**29**	13	11	5	5	16	8	13
E1.L1	**34**	11	13	5	8	8	8	13
E3.L3	0	5	**50**	16	8	21	0	0
E6.L6	0	0	29	11	0	**58**	0	3
E8.L8	8	5	5	5	16	11	16	**34**

Note. I1-I8 denote the eight levels of the imperfectly predictive cue. E1-E8 denote the eight levels of the earlier learned perfectly predictive cue. L1-L8 denote the eight levels of the later learned perfectly predictive cue. A dot (period) separates two cues that appeared together. Numbers in **bold font** indicate the modal response in each row.

$p < .05$. There was also extrapolation of blocker dominance, in that the frequency of responding with extreme (never trained) scale values (”1” or ”2” responses to A8C1 and ”7” or ”8” responses to A1C8) was significantly higher than would be expected by chance responding across the entire scale, $\chi^2_{(df=1,N=68)} = 22.67$, $p < .001$.

Highlighting, positive function group

Table 5 shows complete results. For cues E and L by themselves, there was clear extrapolation of the positive and negative functions, respectively, as highlighted in Table 5 by the bold font for the modal response to each cue. Inferential statistics confirm this observation, in that the frequency of responding ”1” or ”2” to E1, and ”7” or ”8” to E8, was significantly greater than what would be expected from uniform responding across the scale range, $\chi^2_{(df=1,N=50)} = 69.4$, $p < .001$; and, the frequency of responding ”1” or ”2” to L8 and ”7” or ”8” to L1 was significantly greater than uniform responding, $\chi^2_{(df=1,N=50)} = 80.7$, $p < .001$. Other results are discussed in the body of the article.

Highlighting, negative function group

Of 28 participants who experienced the negative function for the early trained cues, 9 failed to reach the learning criterion, yielding N=19 for subsequent analyses. This lower proportion of learners is to be expected because it is well known that learning a negative function is more difficult than learning a positive function (e.g. Brehmer, 1971, 1973). Table 6 shows the results from the test phase.

Performance on the training items was fair in the test phase, averaging 85% correct on the I.E items and 78% correct on the I.L items.

For cues E and L by themselves, there was extrapolation of the negative and positive functions, respectively, as highlighted in Table 6 by the bold font for the modal response to each cue, although the extrapolation was not as sharp for E (the negative function) as for L (the positive function). Inferential statistics confirm this observation, in that the frequency of responding "1" or "2" to E8, and "7" or "8" to E1, was significantly greater than what would be expected from uniform responding across the scale range, $\chi^2_{(df=1,N=38)} = 7.89$, $p < .001$; and, the frequency of responding "1" or "2" to L8 and "7" or "8" to L1 was significantly greater than uniform responding, $\chi^2_{(df=1,N=38)} = 29.51$, $p < .001$.

Most importantly, responses to cue I by itself showed a clear preference for the early-learned outcome, and responses to conflicting cues E.L showed a preference for the later learned outcome. For both types of cues, there was robust extrapolation of these trends beyond the training domain. These trends can be seen clearly in Table 6 by the bold-font modal responses. As observed previously for the positive-function case, responses to the conflicting cues E.L showed a weaker, secondary peak along the function for the early learned outcome.

Inferential statistics support these observations. For cue I by itself, the frequency of "1" or "2" responses to I8 and "7" or "8" responses to I1 was greater than uniform responding across the scale range, $\chi^2_{(df=1,N=76)} = 15.8$, $p < .001$. The frequency of "3" or "4" responses to I6 and "5" or "6" responses to I3 was greater than the frequency of "5" or "6" responses to I6 and "3" or "4" responses to I3, $\chi^2_{(df=1,N=69)} = 17.8$, $p < .001$. For conflicting cues E.L, the frequency of responding along the positive diagonal was significantly greater than the frequency of responding along the negative diagonal, $\chi^2_{(df=1,N=122)} = 15.9$, $p < .001$.