Error-Reduction and Simplicity: Opposing Goals in Classification Learning

Mark Blair (mrblair@indiana.edu)
Indiana University, Department of Psychology
1101 E. Tenth Street, Bloomington, IN 47405 USA

Abstract

Studies of real world experts show that they use different and subtler regularities than novices to make effective classifications. In laboratory studies of learning however, participants have a strong preference for simple cue sets, even at the expense of accuracy. The present experiments investigate participants' ability to use subtle stimulus dimensions in order to eliminate category exceptions. Results show that some participants were able to use the optimal 3-cue set, but many could not. When there were two optimal cue sets, one with 2 dimensions and one with 3, participants favored the simpler set, even though it meant ignoring an obvious and diagnostic cue. Overall there were wide individual differences, with almost every cue set adopted by some participants. Current theories of attention that posit rapid shifts of learned attention offer promise in accounting for the results.

For any organism to adapt successfully it must become sensitive to meaningful regularities in the environment. Humans have developed flexible learning systems, allowing them to rapidly adjust to changing environments. Learning concepts, representations of classes of stimuli that 4require an equivalent response, conserves resources by reducing the amount of information that needs to be processed from the environment, and also allows for generalization to related, novel circumstances. Representing a complex environment, with abundant interdependencies and subtle regularities requires a rich set of concepts.

To perform this function, the human perceptual system can attend selectively, become sensitized to highly complex stimulus dimensions, and even create novel functional features. These processes affect the perception of a stimulus and therefore alter the representation of that stimulus. Differences in cue use and representation across experts and novices appear in many areas such as biology (Boster & Johnson, 1989), physics (Chi, Feltovich, & Glaser, 1981), computer programming (Davies, 1994), wine tasting (Solomon, 1997), bird identification (Johnson & Mervis, 1997) and pocket billiards (Blair & McBeath, 2001). Despite an abundance of differences between expert and novice differences in cue use, laboratory studies of learning (specifically "knowledge restructuring") have shown that participants have had a strong resistance to using new information and more complicated cue sets, even though they would afford better performance (Lewandowsky, Kalish & Griffiths, 2000).

To the extent that learning is error driven, exceptions in the cue set provide a powerful motivator to incorporate new dimensions, however, there is also a pressure toward simplicity. Additional dimensions, which may a space in which the categories separate, can be expensive to represent. Completely altering the dimensions used for categorization can require more energy than using unique stimulus-level elements to classify exceptions. This leads to the memorization of exceptions, rather than a refining of the cue set. It is clear that category learning is influenced by opposing forces; one to enlarge, and one to reduce the dimensionality of the cue set.

The present research examines the complementary processes of the expansion and reduction of the cue set toward effective and efficient representation. The goal of the present study is to verify that participants can and do optimize their cue sets using both expansion and reduction when learning categories with subtle dimensions. Previous research has demonstrated these complimentary processes by manipulating the stimulus set to provide a new dimension (Blair & Homa, 2003b). These studies produced wide individual differences, with many participants incorporating new dimensions to eliminate category exceptions, and many others choosing to rely on simple cue sets which result in many exceptions and significant error. The Blair and Homa (2003b) studies also showed that participants can shift to optimal spaces if they are less complex. Overall, in these studies participants demonstrated both the flexibility found in studies of expertise and the insistence on simplicity found in studies of knowledge restructuring.

Many real world category learning problems do not involve learning new information never experienced before, but rather involve learning to be sensitive to stimulus features that have existed all along, but may have been overlooked for more salient dimensions. For example in bird identification, color is an obvious perceptual cue, used by experts and novices alike. To tell the difference between a Hepatic Tanager and a Summer Tanager, both of which are predominantly red, one must notice the color of the bill and whether the bird has a gray ear patch or not. These are features that novices are prone to miss. The present studies use two obvious dimensions as well as a subtle third dimension as a direct analogy to those cases.

Experiment 1

In Experiment 1, a sequential presentation same-different task was employed to examine the discriminability of the three dimensions used in the remaining experiments, and an additional dimension (color) used in a related set of studies (Blair & Homa, 2003b). If the stimuli are to be used in later experiments, they should be of roughly equal discriminability, with the exception of tail bumpiness, which should be significantly less discriminable than the other three dimensions.

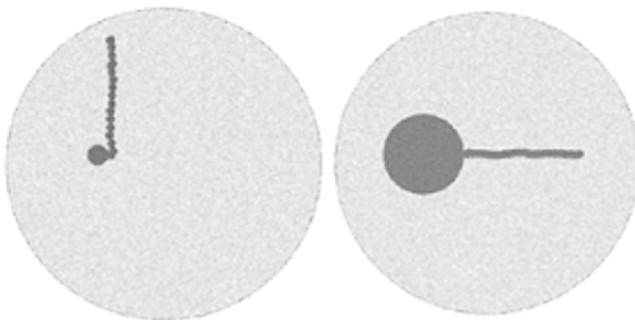

Figure 1: Two Example Stimuli. Tail angle, head diameter, and bumpiness of the tail are the three primary dimensions of variation. Stimuli could take on any one of 11 values on or between the two extremes shown here.

Method

Participants Participants were 26 undergraduates from Arizona State University who participated for course credit. All participants were naïve as to the purpose of the test.

Stimuli Examples of the stimuli are shown in Figure 1. These stimuli were created in the graphics program Adobe Photoshop. Designed to look like something seen under a microscope, these fictitious microorganisms offer a ready analogy to expertise domains such as medical diagnostics. Stimuli varied in one of four different dimensions: head diameter, tail angle, color, and tail bumpiness. Head diameter varied from 25 to 100 pixels. Tail angle varied from 0 to 90 degrees. Color, in RGB values, ranged from 65-75-230 to 139-75-30. These colors were set so that they were of equivalent luminosity, that is, they look the same shade of gray if viewed as a black and white image. The tail bumpiness was created by using brushes with "spacing" set from 60 to 90. The full range of the variation was broken up into 11 equal sized steps, thus dimensions were always one of those 11 values. In addition to possible variations in the four principle dimensions, the tail of each microorganism was hand drawn on top of a line of the correct angle and thus represented a source of stimulus-specific variation. A Gaussian noise filter (30-unit) was also applied to each stimulus. Afterward, a "crystallization" filter (3-pixel) was applied to the stimulus area around, but not including, the microorganism.

Procedure The entire experimental task, including instructions, was displayed on computer. The experimental task was a sequential same-different task in which participants were shown two stimuli, one at a time, and asked to judge whether they were the same or different. Participants were shown two example stimuli and the four main dimensions of variation were indicated. Participants were instructed that only variations on the four consequential dimensions should elicit a 'different' response, any other variations were to be ignored. On each trial, the first stimulus was shown for 2000 msec, and then the screen went blank for 1000 msec. Finally, the second stimulus appeared, and remained on screen until the participant responded. There were 144 trials and of the 144 stimulus pairs, 72 were 'same' pairs and 72 were 'different'. Of the 72 'same' trials, 36 were pairs showing exactly the same stimulus and 36 were pairs of stimuli that had the same values on the consequential dimensions, but were created separately. Of the 72 'different' pairs, there were three stimuli from each of three levels of variation (1, 2 and 4 units) from each of the four main dimensions (head diameter, tail angle, color, and tail bumpiness). The lower, upper and middle parts of the range of variation were used for each trio of stimuli associated with a level of variation. For example, for 1 unit variations on head diameter, the three stimulus pairs might include a pair with 25 and 32-pixel heads, a pair with 93 and 86-pixel heads and a pair with 48 and 54-pixel heads. The values of the remaining three dimensions, which did not vary between members of a stimulus pair, were randomly assigned. They were approximately equally distributed across the possible range of values.

Results and Discussion

To establish that the three primary dimensions are of roughly equal discriminability and that they are all more discriminable than the subtle dimension, a single-factor repeated measures ANOVA was run, using change type (diameter, angle, color, bumpiness, values and same). Results, depicted in Figure 2, showed a significant main effect of change type, $F(5,120)=104.80$, $p<.0001$. Scheffe post-hoc tests revealed that color was not significantly different from either angle or diameter, but that performance on differences in angle were detected significantly more often than differences in diameter, showing that these three dimensions are roughly, but not perfectly equal. The

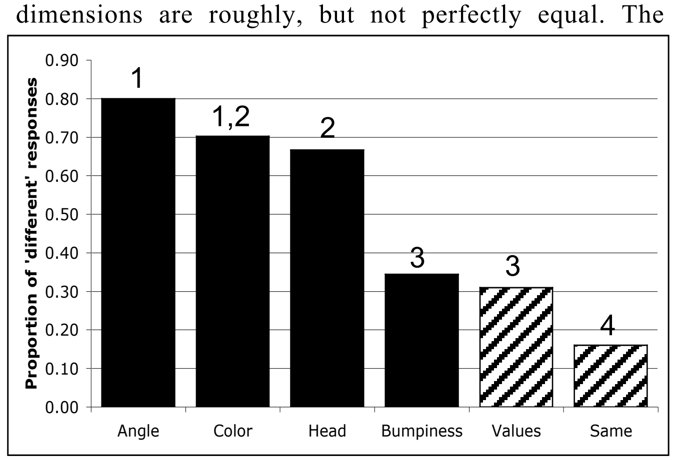

Figure 2: Data from Experiment 1. Groups with different numbers above the bar are significantly different from one another. The striped bars indicate trial blocks where a 'different' was incorrect.

proportions of 'different' responses for each of these three dimensions were significantly higher than for 'same' trials, establishing that participants have the ability to detect these changes. Performance on these three dimensions was also significantly better than on bumpiness, thus, supporting its classification as a subtle dimension. Despite worse performance on the subtle dimension, participants responded 'different' significantly more frequently on trials with changes on this dimension, than on 'same' trials. Finally, performance on 'bumpiness' trials did not differ significantly from performance on 'values' trials, where changes in the way the stimuli were drawn preserved the values of the four key dimensions. Because participants were instructed to ignore the kinds of changes that occur on 'values' trials, they were likely responding without being aware of what kind of changes they saw. The equivalence of performance on these two trial types suggests that participants may be responding with an equivalent lack of awareness on 'bumpiness' trials. This further supports the classification of the bumpiness dimension as subtle.

Experiment 2

Experiment 2 used the subtle dimension of tail bumpiness. This procedure makes this experiment an apt analog to real world categories that experts must master. The effect of pointing out the subtle dimension to participants was investigated using three between-subjects conditions: Help, No-Help (NH) and No-Help/Help (NH-H). In the Help condition participants were told in the instructions at the beginning of the experiment that tail bumpiness is an important cue in helping to classify the stimuli correctly. In the NH condition, participants were never told about tail bumpiness. In the NH-H condition, participants were given this information at the beginning of Stage 2.

Method

Participants Participants were 83 Arizona State University undergraduates enrolled in an introductory psychology course. They participated for course credit. They were randomly assigned to one of three conditions: Help (n=33), NH (n=24), NH-H (n=26).

Stimuli The stimuli had three principal dimensions of variation: head size, tail angle and tail bumpiness as described in Experiment 1.

Structure The two categories used were linearly separable only in three dimensions; therefore the best lower dimensional bounds always had exceptions. The best two-dimensional linear decision boundary left 14% of the exemplars as exceptions and the best single dimension linear decision boundary left 30% as exceptions. Figure 3 shows the training stimuli plotted in the angle/diameter space and also bumpiness/angle/diameter space.

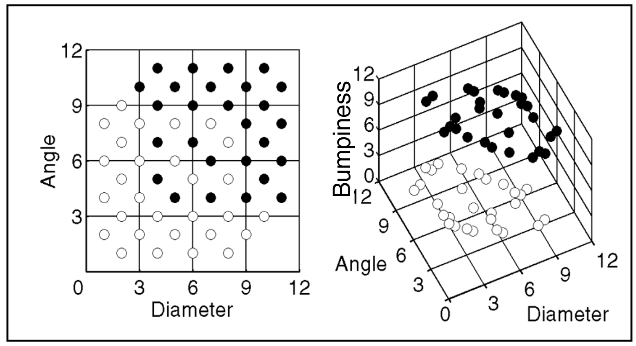

Figure 3: The stimulus space used in Experiment 2, plotted for stimuli with and without the subtle bumpiness dimension.

Procedure The experiment had two stages with each stage consisting of a learning phase and a transfer phase. The learning phase included four blocks of 56 trials. Within a trial block every stimulus in the learning set was presented once. The presentation order was randomized for each trial block, and for each participant. On each trial, the participant was shown a stimulus and asked to classify it as 'normal' or 'deviant'. Once the participants indicated their choice by pressing the appropriate key ('n' or 'd') on the keyboard, the correct answer was presented next to the stimulus for 2000 msec. If the participant answered incorrectly, the feedback was red instead of black. After the learning phase, participants performed a transfer task in which they classified two cycles of 16 novel stimuli. The transfer set presented after the three-dimensional stimulus set consisted of 8 stimuli on either side of the best three-dimensional linear decision boundary at a range of values. These values were such that if participants were using the best combination of angle and diameter dimensions, or any one of the three dimensions alone, they would achieve 50% correct. Participants were not given feedback during the transfer task. Stage 2 was a repetition of Stage 1, except for the instructions, as determined by the condition. For all conditions, the instructions encouraged participants to achieve perfect classification for both Stage 1 and Stage 2.

Results and Discussion

Overall performance was equally good whether or not participants were told about the usefulness of the subtle dimension (Help, M=87%; NH, M=88%; NH-H, M=88%). The dimension of primary interest is tail bumpiness. Single group t-tests against zero revealed that NH-H condition showed significant increase in the use of tail bumpiness in the second stage, $t(25)=2.34$, $p<.05$, but the other conditions showed no significant change in tail bumpiness from Stage 1 to Stage 2. The instructions clearly increased the use of tail bumpiness. The NH group had the lowest use of tail bumpiness, with the NH-H group significantly increasing their use of tail bumpiness, after being instructed to do so.

Table 1: Percentages of participants for which each cue set accounts for most of the transfer test responses from Experiment 2.

Condition	Stage	Bump	Angle	Diam	D/A	B/D/A
NH	1	36%	11%	18%	29%	7%
NH-H	1	29%	14%	14%	32%	11%
H	1	43%	14%	6%	26%	11%
NH	2	16%	32%	13%	26%	13%
NH-H	2	56%	11%	11%	4%	19%
H	2	32%	22%	7%	15%	24%

Individual differences in the adoption of the various cue sets were assessed by the transfer test. Participants were sorted according to the cue set that best matched their responses. Because all dimensions were present in the transfer stimuli, it was possible for a participant's responses to fit two different cue sets equally well. In such cases both cue sets were counted and the percentage reported was calculated across all preferences not all participants. That said, the large majority of participants preferred only one cue set. The percentages of preferences for each cue set are reported in Table 1.

Overall, many participants seemed to make use of the subtle dimension of tail bumpiness early in training; even without having it brought to their attention. Use of the subtle dimension was increased by instructions however, and though participants in the NH condition showed no performance deficit by the end of Stage 2, the transfer test reveals that they used tail bumpiness least. The general findings of this experiment mirror the Blair & Homa (2003b) studies using an obvious 3rd dimension, instead of the subtle one, namely some participants incorporated the 3rd dimension while others stuck with one or two sub-optimal dimensions. One question still open is whether or not more participants would adopt a cue set which separates the categories if it had fewer dimensions. This question is addressed in Experiment 3.

Experiment 3

In Experiment 2, it was shown that many participants rapidly detected and used diagnostic information, even though it was subtle. Some participants can readily use a three-dimensional cue set if the categories are made separable. For Experiment 3, a new stimulus space was created by changing the values on the subtle dimension. Like the space in Experiment 2, this space was not linearly separable using the two regular dimensions (head diameter and tail angle) but was separable by also considering the subtle dimension (tail bumpiness). Unlike the space for Experiment 2, this space was also separable when considering only head diameter and tail bumpiness. The separable 2-D and 3-D spaces are shown in Figure 4. These categories allow participants to collapse their cue set to only two dimensions with no loss of accuracy. The primary objective of this experiment was to assess the degree to which participants are able find the most efficient cue set if it required using fewer dimensions.

The instruction manipulation was dropped for this experiment; all participants were in the equivalent of the NH condition. Also, individual differences in this experiment were expected to be more important because there were multiple effective strategies. Accordingly, a larger number of participants were tested.

Method

Participants Participants were 96 Arizona State University undergraduates enrolled in an introductory psychology course. They participated in the experiment to fulfill a course requirement.

Structure In order to detect the use of the various possible 2-D spaces during transfer, the stimulus values for the training and transfer stimulus spaces were altered from Experiment 2. As in the Experiment 2 space, individually, diameter and angle were 70% predictive. The bumpiness dimension was 73% predictive. In 2-D space (head diameter and tail angle) this stimulus space is identical to the Experiment 2 space, that is, 8 of the 56 stimuli (14%) were exceptions. In 3-D space, the two categories were linearly separable. Also, in the 2-D space defined by head diameter and tail bumpiness the categories were linearly separable. The two spaces where the categories are linearly separable are shown in Figure 4. The category structure (mean within-category distance divided by mean between-category distance) was .60 for the categories represented in the diameter/bumpiness space and .66 for the categories when represented in the 3-D space.

Procedure The procedure was identical to Experiment 2 except the transfer tasks, which involved 2 cycles through a transfer set with 12 stimuli.

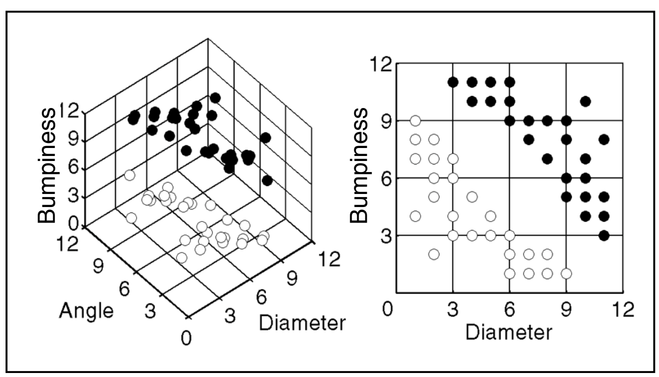

Figure 4: This stimulus space used for Experiment 3. This space was modified from the space for Experiment 2 so that the categories are separable in both the B/A/D space (plotted on the left) and the B/D space (plotted on the right).

Results and Discussion

Overall, performance improved across trial blocks and participants averaged 83% correct on the final trial block. In the transfer task, the measures of the three single dimension cues (bumpiness, angle and head diameter) and three two dimension cue sets (bumpiness/angle, bumpiness/diameter and diameter/angle) are not independent and require some explanation. Dimension use, as in Experiment 2, is measured as the proportion of trials that would be supported by that dimension or pair of dimensions. For each of the six cues, 8 of the 12 transfer trials present a stimulus on which the values of the cues suggest one category over another, and 4 of the 12 transfer trials present a stimulus for which the cue is neutral. Each single dimension and its dual-dimension opposite (e.g., diameter and bumpiness/angle) use the identical eight trials, and so are perfectly negatively correlated. This means that if a participant chose according to the value of tail bumpiness on .88 proportion of the trials, then they must also have chosen according to head/angle on the remaining .12 proportion of the trials. The other cues are only partially correlated. The overall attention use is forced to sum to 1.5 for the three single dimension cues and to 3.0 for all six dimensions. As an example, a participant might have a balanced attentional profile, equally distributing attention to all dimensions in combination (the 3-D solution). This profile would conform to each cue set on .50 proportion of the trials. A different participant might not use the subtle tail bumpiness cue at all. This profile would yield a 1.0 proportion for D/A (and 0 for bumpiness) and, assuming equal weight to the remaining dimensions, .25 for B/A and B/D (and therefore .75 for diameter and angle individually).

In the first transfer task, the use of the head dimension is consistent with the highest proportion of trials (.63) with D/A (.61) and B/D (.52) the next two highest. The emphasis on these three dimensions increased after the second stage of training (.70, .62 and .58 respectively). Both the increase in head and the increase in B/D (or, alternatively, the decrease in use of angle) were significantly different than chance, $t(95)=2.89$, $p<.01$ and $t(95)=3.20$, $p<.01$. Participants seem to increasingly ignore the information given by angle in favor of reliance on head diameter, even though angle provides easier discriminations (Experiment 1).

As in Experiment 2, a table summarizing individual data was created. The number of participants for which a cue or cue set received the highest use during the final transfer test was recorded. Twelve participants had equal endorsement for two different cue sets. In these cases, each cue was given credit, so the total number of endorsements is 108, even though there were only 96 participants. The percentage of the total number of preferences is shown for each cue in Table 2.

Table 2: Percentages of participants for which each cue set accounts for most of the transfer test responses from Experiment 3.

Stage	Bump	Angle	Diam	D/A	B/D	B/A	B/D/A
1	6%	16%	29%	24%	17%	4%	5%
2	1%	6%	38%	23%	23%	1%	8%

This tabulation reveals that 30% of the participants used a space in which the categories were separable, with nearly three times as many preferring the simpler B/D space to the B/D/A space by the end of training. This more than the 13% in the equivalent condition from Experiment 2. The reduction in angle use that is equivalent to the adoption of the B/D cue set is made more dramatic in that the angle dimension yielded the most accurate discriminations in Experiment 1. Further, both tail angle and tail bumpiness are features of the same component of the stimulus. On the other hand, there were still two thirds of the participants who used only obvious dimensions and failed to adopt a cue set which separates the categories. The preference for simplicity may extend to the discriminability of the cues as well as their number.

General Discussion

The experiments reported here investigated participants' ability to use subtle perceptual cues to disambiguate overlapping categories. In Experiment 1, it was established that the primary dimensions of variation, head diameter and tail angle, were roughly equated for discriminability, and the subtle dimension of tail bumpiness was much less discriminable. Experiment 2 demonstrated that some, but not all participants were able to use the subtle dimension in conjunction with the other dimensions to eliminate exceptions to the obvious dimensions. Experiment 3 showed that participants favored using a 2-cue optimal cue set over a 3-cue one.

There are two aspects of the present data that are challenging from a modeling perspective. The first is the widespread disregard for diagnostic cues. In some participants, this shows up as an inability to use the bumpiness cue or perhaps even a total reliance on just one obvious cue. These expedient cue sets could not yield enough information to result in perfect performance without additional memorization of exceptions. Even for participants who chose one of the cue sets that separated the categories, there was a strong preference for the simpler (2-d) set. Use of the angle dimension, which was part of the optimal 3d cue set, and which was independently diagnostic, decreased with training. The second aspect of the data that is challenging is the broad individual differences. Nearly every possible cue set was used in all phases of the reported experiments. Category learning models which adjust attention weights based on gradient descent on error (i.e., the backprop algorithm) will tend to converge on a set of optimal weights, rather than showing dramatic differences in predicted performance (see discussion in Kruschke, 2001). These optimal weights will also be positive for any informative dimensions. More recent models of attention in associative learning that incorporate rapid shifts of attention in conjunction with annealed learning rates show more promise of fitting our data. The rapid shifts of attention can lead to shifts away from dimensions before associations can form, and the annealed learning results in a progressive discounting of error, leading to participants getting frozen in a sub-optimal space. Kruschke and Johansen's (1999) RASHNL model has fit similar data in probabilistic category learning task.

The present study is related to recent work on human concept learning that is based on a simplicity principle (Feldman, 2003). This principle suggests that the ease with which categories can be learned is related to how incompressible or complex the categories are, that is, the length of its minimum description. Maximally complex categories, for example one consisting of a mailman, a speedboat, and a jelly doughnut, have no regularities at all, and the minimum description is simply a list of its members. Simpler categories, for example: big blue triangle, big red circle, and big yellow square, can be compressed down to exclude some of the data from the examples, leaving a smaller description, in this case "big" things. In the context of the current work, it could be said that incorporating the information from the bumpiness dimension, while increasing the number of dimensions, decreases the overall category description length, because the exceptions do not have to be explicitly encoded. Participants in Experiment 2, some of who showed increasing use of the bumpiness dimension, but also decreasing use of the angle dimension, also seemed to use a simplicity principle. Angle use decreased not because it was uninformative, but because it was not part of the minimal description. In addition to reworking their description of their overall category regularities, participants can and do augment their category representation by identifying and memorizing specific exception stimuli. In other words, simply adding any exceptions to the representation they have already formed. Several results, including the present data, suggest that this strategy is not uncommon in some typical category learning paradigms (Blair & Homa, 2001). This focus on individuals rather than on category level regularities seems to occur even in some separable categories if they are small and weakly structured (Blair & Homa, 2003). These results highlight the potential disparity between mathematical complexity and psychological complexity, and emphasize the importance of understanding how the cognitive system implements complexity minimization. A precise account of the relative costs of adding or shifting dimensions versus memorizing exceptions will certainly involve a better understanding of how attentional, perceptual, and memorial processes interact as classification expertise develops.

Acknowledgments

This work was completed in partial fulfillment of the Ph.D. requirements in psychology at Arizona State University. I would like to thank Steve Goldinger, Mike McBeath and Sue Somerville who comprised my excellent dissertation committee, and especially Don Homa, my mentor and committee chair. I would also like to thank Jacqueline Blair for helpful comments on an earlier draft of this paper and Chris Summers for help in conducting these experiments.

References

Blair, M. & Homa, D. (2001). Expanding the search for a linear separability constraint on category learning. *Memory & Cognition 29*. 1153-1164.

Blair, M. & Homa, D. (2003). As easy to memorize as they are to categorize: The 5-4 categories and the category advantage. *Memory & Cognition 31*. 1293-1301.

Blair, M. & Homa, D. (2003b). Integrating Novel Dimensions to Eliminate Category Exceptions: When More Is Less. *Manuscript submitted for publication.*

Blair, M. & McBeath, M. K. (2002). Bangers in billiards - expert/novice differences in shot force. *Manuscript in preparation.*

Boster, J. S., & Johnson, J. C. (1989). Form or function: A comparison of expert and novice judgments of similarity among fish. *American Anthropologist, 91*(4), 866-889.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121-152.

Davies, S. P. (1994). Knowledge restructuring and the acquisition of programming expertise. *International Journal of Human Computer Studies, 40*(4), 703-726.

Feldman, J. (2003). The simplicity principle in human concept learning. *Current Directions in Psychological Science*, 12, 227-232.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology-General, 126*(3), 248-277.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45*(6), 812-863.

Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(5), 1083-1119.

Lewandowsky, S., Kalish, M., & Griffiths, T. L. (2000). Competing strategies in categorization: Expediency and resistance to knowledge restructuring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1666-1684.

Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology, 32*(1), 49-96.

Solomon, G. E. A. (1997). Conceptual change and wine expertise. *Journal of the Learning Sciences, 6*(1), 41-60.