The Influence of Perceptual Difficulty on Family Resemblance Sorting

Fraser Milton (f.n.milton@exeter.ac.uk)
School of Psychology, University of Exeter,
United Kingdom

A.J. Wills (a.j.wills@exeter.ac.uk)
School of Psychology, University of Exeter,
United Kingdom

Abstract

The impact of perceptual difficulty on the prevalence of family resemblance sorting was investigated in a free classification experiment. There were two between-subject conditions, high perceptual difficulty and low perceptual difficulty and participants were asked to sort the stimuli into two groups in the way that seemed most appropriate to them. The results showed that participants in the low perceptual difficulty condition sorted by family resemblance to a greater extent than participants in the high perceptual difficulty condition. This finding extends the work of Milton and Wills (2004), who showed that the level of spatial integration in stimuli is an important determinant of sorting behavior, but whose results were inconclusive on the issue of perceptual difficulty. The current experiment adds to the growing body of work which demonstrates that free sorting behavior can be influenced in a variety of different ways.

Keywords: free classification; family resemblance; unidimensional; perceptual difficulty; match-to-standards.

Introduction

Categorization is a fundamental cognitive mechanism that enables us to function effectively in our everyday environment. However, in view of the immense number of objects we encounter, this process must necessarily be highly constrained. To illustrate this, just ten objects can be partitioned into more than 100,000 separate ways (Ahn & Medin, 1992). A greater knowledge of how we acquire the categories that we have is therefore an important requisite for our understanding of human cognition.

Traditionally, categorization experiments have used a *supervised learning* procedure. Under this procedure, participants are given perfect, trial-by-trial feedback as they learn to predict category membership from the attributes of various exemplars. Whilst such an approach has greatly advanced our understanding of categorization, it tells us little about how people construct categories in situations where feedback is rarer, non-trial specific, or absent. This question is important because, outside the laboratory, people rarely receive the level of feedback that occurs in supervised categorization experiments. Free classification (e.g. Imai & Garner, 1965) - also known as free sorting (e.g. Bersted, Brown & Evans, 1969), category construction (e.g. Medin, Wattenmaker & Hampson, 1987) or spontaneous categorization (e.g. Pothos & Chater, 2002) - is a methodology for investigating categorization preferences in the absence of feedback. Participants are given a set of stimuli and asked to sort them in the way that seems most sensible and natural to them. No feedback is given.

One reasonable assumption is that the categories we prefer to create would reflect the underlying structure of objects we encounter outside the laboratory. Perhaps the most influential theory of natural categories is the idea that they are organized around a "family resemblance" structure (e.g. Rosch & Mervis, 1975; Wittgenstein, 1958), in which categories possess a number of characteristic but not defining features. Under a family resemblance structure, an object does not have to possess any particular feature but can be considered a member of that category if it possesses enough characteristic features. Items in a family resemblance structure are organized around overall similarity relations, which maximize within-group similarities and minimize between-group similarities.

Despite the plausibility of a family resemblance theory of natural categories, previous work has shown that when people are asked to free classify stimuli they find it far from natural to sort by family resemblance. In fact, people have a strong tendency to free classify unidimensionally (i.e. on the basis of a single feature, e.g. Ahn & Medin, 1992; Ashby, Queller, & Berretty, 1999; Medin et al., 1987). Whilst manipulations of the method of stimulus presentation (Regehr & Brooks, 1995), the level of spatial integration of the stimuli (Milton & Wills, 2004), time pressure (Milton, Longmore, & Wills, in press), cognitive load (Milton et al., in press), inductive inference (Lassaline & Murphy, 1996), and background knowledge (Spalding & Murphy, 1996) influence the extent of family resemblance sorting, such sorting is still far from ubiquitous. It therefore appears important to understand why the categories we prefer to create do not reflect the commonly assumed underlying structure of natural world categories.

The starting point for the current study was the finding of Milton and Wills (2004) that the physical characteristics of the stimuli have an important influence on the prevalence of family resemblance sorting. In two experiments (Experiments 4 and 5), Milton and Wills investigated factorially the influence that the level of spatial integration and the level of perceptual difficulty have on sorting behavior (Figure 1 shows the prototypes of the stimuli used in Experiment 4). They showed that stimuli that were more spatially separable resulted in significantly more family resemblance sorting than stimuli that were more spatially integrated. In contrast, whilst there was a trend in both experiments for the stimuli of low perceptual difficulty to evoke family resemblance sorting to a greater extent than the stimuli of high perceptual difficulty, in neither experiment did this trend approach significance. Perceptual difficulty interacted with the level of spatial integration in Experiment 4, but this effect was not replicated in Experiment 5.

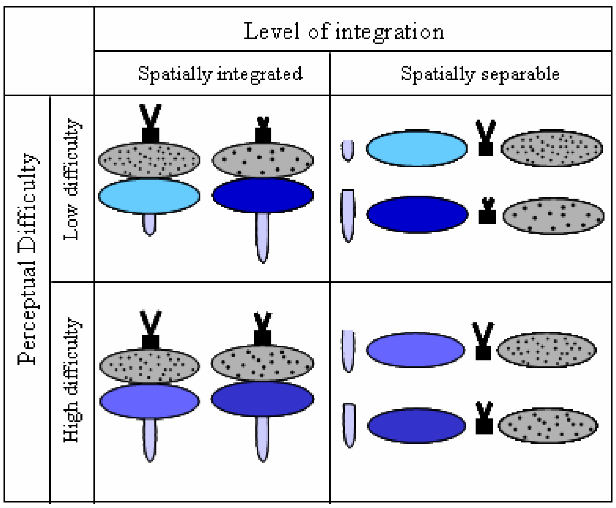

Figure 1: The prototypes for the four stimulus sets used in Experiment 4 of Milton and Wills (2004).

The finding that the more spatially separable stimuli produced a greater level of family resemblance sorting is somewhat surprising when one considers previous work, which has shown that integral stimuli evoke more family resemblance sorting than separable stimuli (e.g. Handel & Imai, 1972; Kemler & L.B. Smith, 1979). Whilst none of the stimuli Milton and Wills (2004) used are likely to be integral according to Garner's (1974) definition, integrality / separability has often been considered to be a continuum rather than a dichotomy (e.g. L.B. Smith & Kemler, 1978) and it seems reasonable to argue that the spatially separable stimuli are more separable than the spatially integrated stimuli. According to this line of reasoning, it might therefore have been expected that the spatially integrated stimuli would evoke more family resemblance sorting. In fact, the opposite occurred.

The explanation that Milton and Wills (2004) proposed to explain their somewhat surprising results was that the family resemblance sorting they observed was the result of an analytic rather than a non-analytic processing strategy. Typically, an analytic strategy has been regarded as a relatively effortful strategy in which the stimulus is broken down into its constituent dimensions before a decision is made whereas a non-analytic strategy is a primitive strategy in which dimensions are perceived as a unitary whole (e.g. Kemler Nelson, 1984). It has often been thought that family resemblance sorting is the result of a non-analytic strategy, whilst unidimensional sorting is due to an analytic strategy (e.g. Kemler Nelson, 1984; J.D. Smith & Kemler Nelson, 1984; Ward, 1983). However, it is unclear how a non-analytic account of family resemblance sorting can account for the spatial integration effect that Milton and Wills (2004) observed.

Milton and Wills (2004) proposed that, under certain conditions, family resemblance sorting can be the result of a dimensional summation strategy in which participants break down the stimulus into its constituent dimensions and place the stimulus into the category for which it has more characteristic features. This account assumes that both family resemblance and unidimensional sorting can be the result of an analytic process, with family resemblance sorting simply being a more complex and time-consuming analytic process than unidimensional sorting. It seems plausible to assume that spatially separating the stimulus dimensions makes it easier to differentiate those dimensions. This, in turn, seems likely to make a dimensional summation strategy less effortful and quicker to apply for the spatially separable stimuli than for the spatially integrated stimuli. This might be anticipated to increase family resemblance sorting, which is what Milton and Wills (2004) found.

The idea that, under certain conditions, family resemblance sorting can be the result of an effortful, analytic, dimensional summation strategy has received some support in a series of studies conducted by Milton et al. (in press). In particular, they showed that participants under moderate time pressure produced significantly fewer family resemblance sorts than participants under very little time pressure. Milton et al. (in press) also showed that participants under a moderate concurrent cognitive load produced significantly fewer family resemblance sorts than participants under no concurrent load. Such findings are in accord with the idea that family resemblance sorting can be a time consuming strategy that requires more cognitive resources than unidimensional sorting.

Whilst this subsequent work provides some evidence for the idea that family resemblance sorting can be the result of an effortful analytic strategy, it provides little further insight into the role of stimulus differentiation assumed by Milton and Wills (2004). Indeed, whilst a differentiation account provides an explanation of the spatial integration effect, it is unclear why, according to this account, Milton and Wills (2004) failed to find any reliable effect of perceptual difficulty. One might have anticipated that stimuli of low perceptual difficulty (ie. easily discriminable stimuli) would evoke a greater level of family resemblance sorting than stimuli of greater perceptual difficulty because it would be easier to differentiate the dimensions when there was a greater difference between the values. For this reason, it appears important to further investigate the effect that the level of perceptual difficulty has on the prevalence of family resemblance sorting if one is to better understand the influence that differentiation has on sorting behavior.

One explanation for the failure of Milton and Wills (2004) to show a significant effect of perceptual difficulty is that the dimensions may not have been manipulated over a sufficiently wide set of values. It is possible that if the values of the dimensions were differentiated more strongly than was the case in Milton and Wills (2004) then this could produce a change in strategy similar to that observed for the spatial integration variable. This hypothesis was investigated in the current experiment.

Method

Participants

Students from the University of Exeter participated either for course credits or for a payment of £2. There were twenty-four participants (12 in each condition) [1] who were tested individually in a quiet testing cubicle.

Stimuli

The stimuli had the same abstract structure as employed by Medin et al. (1987). This stimulus structure is shown in Table 1. The stimulus set consisted of four binary-valued dimensions (D1-D4) and the stimuli were organized around two prototypes, each representative of one of the categories. These prototypes were constructed by taking all the positive values on the dimensions for one of the stimuli (1, 1, 1, 1) and all of the zero values on the dimensions (0, 0, 0, 0) for the other category. The rest of the stimuli (one-aways) were mild distortions of the two prototypes in that they had three features characteristic of their category and one atypical feature more characteristic of the other category. In total there were 10 stimuli in the set. Sorting the stimuli by family resemblance, as shown in Table 1, maximizes within-group similarities and minimizes between-group similarities.

Table 1: Abstract stimulus structure

Category A				Category B			
D1	D2	D3	D4	D1	D2	D3	D4
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	1
1	1	0	1	0	0	1	0
1	0	1	1	0	1	0	0
0	1	1	1	1	0	0	0

Note. Each row (within each category) describes a different stimulus. D = dimension: 1 and 0 represent the values of each dimension.

[1] The perceptually difficult condition has previously been reported (Milton & Wills, 2004, Experiment 4). The perceptually simple condition has not been previously reported. The population was the same for both conditions and the procedures were identical.

Two different stimulus sets were used – a low perceptual difficulty set and a high perceptual difficulty set. These stimulus sets were identical except that the difference between the values the dimensions took was greater for the stimuli of low perceptual difficulty. This should make it easier to differentiate the dimensions in the low perceptual difficulty condition.

The stimuli were based on the artificial butterflies originally used in Experiment 4 of Milton and Wills (2004). The stimuli varied on four dimensions, the size of the "antennae" (long / short), the number of dots on the top "wings" (many / few), the hue of the bottom "wings" (light blue / dark blue) and the size of the "tail" (short / long). The prototypes of the two stimulus sets are shown in Figure 2.

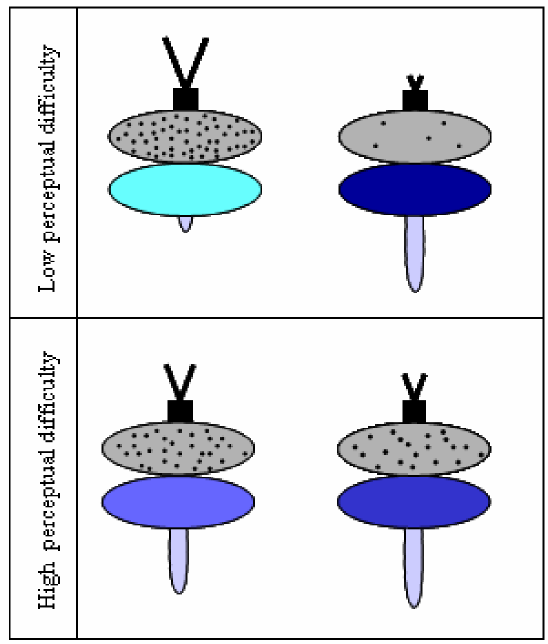

Figure 2: The prototypes of the perceptually simple and perceptually difficult stimuli.

Procedure

The procedure for the two conditions (perceptually simple and perceptually difficult) was identical except for the difference in stimuli. Participants were introduced to the stimuli by the pre-sort matching pairs procedure developed by Milton and Wills (2004). In this procedure, twenty cards were spread out randomly in an array in front of the participant. These cards consisted of the ten stimuli in the set and an identical copy of each of them. The participant had to match these twenty stimuli into the ten identical pairs correctly without feedback. When the participant felt that the pairs had been matched correctly, the pairs were examined, and if there were any errors, participants were asked to match the pairs again. The purpose of this task was to ensure that the participant could fully distinguish the four feature-pairs, because if the participant had not identified all of the features, a family resemblance sort might be difficult to carry out.

After this pre-sort procedure, the method of stimulus presentation for the categorization task was a slight variation on Regehr and Brooks's (1995) match-to-standards procedure and identical to that used by Milton and Wills (2004) and Milton et al. (in press). The two prototypes (presented on cards sized 14.2 cm wide by 8.3 cm high) were placed side-by-side on a table throughout the experiment and participants were informed that these were characteristic of category A and category B (which category each prototype represented was randomized). Participants were then given the ten stimuli in the set (on ten 14.2 cm x 8.3 cm cards) and were asked to look at each stimulus in turn and put it into the group for which they felt it was most representative. Once they had made their decision about a card, participants placed it face down directly below the prototype of the category they felt it most resembled.

Participants were told that there were many ways in which the stimuli could be split and that there was no one correct answer. They were also told that the two groups did not have to be of equal sizes and the only constraints were that each stimulus had to be placed into one of the two groups and that they were not allowed to look through the cards that remained to be sorted or to change previous responses. Participants were informed that there was no time limit and they were encouraged to take as much time as they needed to complete the task to their satisfaction. Once they had finished classifying the stimuli, the participants were asked to explain as precisely as possible the way in which they had classified the ten items.

Analysis of results

Each participant was classified as having produced one of the sort types described below. These sort types are very similar to those employed by Regehr and Brooks (1995) and are identical to those employed by Milton and Wills (2004).

A *unidimensional* or single dimensional sort is based on a single dimension of the stimulus. It does not matter which dimension is used as the basis of sorting, so long as all of the positive values for the chosen dimension are in one category and all of the zero values for that dimension are in the other category. Additionally, in order to receive this classification, the participant has to describe their sort as being based on a single dimension.

Participants were considered to have produced a *one-away unidimensional* sort if they described their sorting as being driven by a single dimension but there was a solitary error in their classification. This means that nine of the items were classified on the basis of a single dimension but the other item was placed into the wrong category.

A *family resemblance sort*, also commonly known as an "*overall similarity*" sort, has a structure identical to that shown in Table 1. In order to receive this classification, the participant has to place each of the prototypes, along with their derived one-aways, into separate categories without error. Additionally, they have to describe their strategy as being based either on general similarity or on placing each item into the category with which it had more features in common.

A *one-away family resemblance* sort is similar to the one-away unidimensional sort with the exception that the error occurred in a sort that was otherwise family resemblance.

Any classifications produced by a participant other than those described above were classified as *other* sorts, even if the description given by the participant fitted one of the sorts described above.

Results

An independent samples t-test was conducted to compare the amount of time that it took participants in the perceptually simple condition to complete the pre-sort matching pairs task compared to those in the perceptually difficult condition. This analysis revealed that participants in the perceptually simple condition (M = 62.89s, SD = 21.42s) took significantly less time to complete the pre-sort task than those in the perceptually difficult condition (M = 431.80s, SD = 165.67s), $t\ (22) = 7.65$, $p < .001$ (the Welch-Satterthwaite correction for unequal variances was applied). This finding confirms that participants found it easier to differentiate the dimensions for the perceptually simple stimulus set than they did for the perceptually difficult set.

The effect of perceptual difficulty on family resemblance and unidimensional sorting is shown in Table 2. As in Milton and Wills (2004) and Milton et al. (in press), the family resemblance and unidmensional categories were combined with their respective 1-aways to produce sum unidimensional and sum family resemblance categories. A chi-square analysis (using the sum FR and sum UD categories) demonstrated that there was a significant effect of perceptual difficulty on sort type, χ^2 (1, N = 24) = 6.33, $p < .02$, [2] with participants in the perceptually simple condition producing a significantly greater ratio of family resemblance to

[2] Yates's correction for 2 x 2 chi-square tables has not been applied. Research has shown that the conventional chi-square for 2 x 2 designs, without correction for continuity, is sufficient to prevent Type I errors (Overall, 1980). Also, the assumption of fixed marginals made by Yates's correction was not applicable for this data set. Additionally, no corrections have been applied for the low expected frequencies of some of the cells. It has been found that small expected frequencies do not increase the chance of type I errors (Bradley, Bradley, McGrath & Cutcomb, 1979). A general discussion of these issues can be found in Howell (2002, pp. 151-152).

Table 2: Sort strategy frequencies

	Sort strategy						
	UD sort			FR sort			Miscellaneous
	1-a UD	UD	Total UD	1-a FR	FR	Total FR	Other
Perceptually difficult	1	11	12	0	0	0	0
Perceptually simple	2	5	7	1	4	5	0

Note. *1-a* = one-away; UD = unidimensional; FR = family resemblance.

unidimensional sorting than those in the perceptually difficult condition. Due to the low sample sizes for the various sort types, we did not analyze the reaction time data for the categorization phase of the experiment.

General Discussion

Milton and Wills (2004) suggested that, under certain conditions, both family resemblance and unidimensional sorting can be the result of an analytic process, with family resemblance sorting simply being a more complex and time-consuming analytic process than unidimensional sorting. It seems plausible to assume that using perceptually simple dimensions makes it easier to differentiate those dimensions. This, in turn, seems likely to make an analytic family resemblance strategy less effortful and quicker to apply for perceptually simple stimuli than for perceptually difficult stimuli.

Given this prediction, it was perhaps somewhat concerning that Milton and Wills (2004) were unable to find any significant effects of perceptual difficulty across two experiments. In the current paper, we hypothesized that this might have simply been because the two perceptual difficulty conditions in Milton and Wills (2004) were not different enough to demonstrate a significant effect with the sample sizes employed. In the current experiment, we increased the difference between the two perceptual difficulty conditions substantially, and found a significant effect of perceptual difficulty.

The current results support Milton and Wills's (2004) theory in the sense they demonstrate the presence of a phenomenon that this theory predicts. We do not wish to claim, however, that this is a *unique* prediction of the Milton and Wills (2004) theory – it seems likely that it would be possible to develop other explanations of the perceptual difficulty phenomenon reported in this paper. Indeed, we are very keen for such theoretical development to take place. However, any such account would also need to explain the growing set of findings of which this is just the latest part. To summarize, it is not only perceptual simplicity that encourages family resemblance sorting in the current procedures; family resemblance sorting is also encouraged by (a) spatial separateness of dimensions (Milton and Wills, 2004), (b) moderate time pressure (Milton, 2006; Milton et al., in press), and (c) concurrent cognitive load (Milton et al., in press).

One prediction derivable from the current findings is that pre-exposure, which has previously been shown to elicit perceptual learning in a free classification paradigm (e.g. Wills & McLaren, 1998), would similarly increase stimulus differentiation. This could make an analytic, family resemblance strategy easier and more likely to be applied for stimuli that have been pre-exposed compared to stimuli where no exposure has been applied. This appears potentially important when one considers that previous free classification studies have generally provided participants with little or no exposure to the stimuli prior to classification. In contrast, we generally have considerable experience with the items that we encounter in our real-world environment. If pre-exposure does facilitate the formation of family resemblance categories then this could provide some explanation for the dearth of such sorting in many previous studies of free classification. This is a question that we are currently investigating.

In conclusion, the current study provides evidence that the level of perceptual difficulty can have a significant effect on the prevalence of family resemblance sorting. As such, this study adds to the growing body of work that demonstrates the subtle factors that can influence the way people create categories in the absence of feedback. In this sense our study adds to previous work that demonstrates that stimulus presentation technique (Regehr & Brooks, 1995), background knowledge (e.g. Spalding & Murphy, 1996), and inductive inference (Lassaline & Murphy, 1996) all have a significant influence on the prevalence of family resemblance sorting.

Acknowledgments

This research was supported by ESRC grants PTA-030-2003-00287 and PTA-026-27-1256 awarded to the first author, and ESRC grant RES-000-22-1779, EC Framework 6 project Grant 516542 (NEST) and Biotechnology and Biological Sciences Research Council Grant 9/S17109 awarded to the second author. The authors thank Lee Brooks for his suggestions and Anna Bullimore for help in collecting the data.

References

Ahn, W.-K., & Medin, D. L. (1992). A two-stage model of category construction. *Cognitive Science, 16*, 81-121.

Ashby, F.G., Queller, S. & Berretty, P.M. (1999) On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics, 61*, 1178-1199.

Bersted, C.T., Brown, B.R., & Evans, S.H. (1969). Free sorting with stimuli in a multidimensional space. *Perception & Psychophysics, 6b*, 409-413.

Bradley, D.R., Bradley, T.D., McGrath, S.E. & Cutcomb, S.D. (1979). Type I error rate of the chi-square test of independence in R x C tables that have small expected frequencies. *Psychological Bulletin, 86*, 1290-1297.

Garner, W.R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Handel, S., & Imai, S. (1972). The free classification of analyzable and unanalyzable stimuli. *Perception and Psychophysics, 12*, 108-116.

Howell, D.C. (2002). *Statistical methods for psychology.* 5th Edition*.* Duxbury Press.

Imai, S. & Garner, W.R. (1965). Discriminability and preference for attributes in free and constrained classification. *Journal of Experimental Psychology, 69*, 596-608.

Kemler, D.G., & Smith, L.B. (1979). Accessing similarity and dimensional relations: Effects of integrality and separability on the discovery of complex concepts. *Journal of Experimental Psychology: General, 108*, 133-150.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior, 23,* 734-759.

Lassaline, M.E. & Murphy, G.L. (1996). Induction and category coherence. *Psychonomic Bulletin & Review, 3*, 95-99.

Medin, D.L., Wattenmaker, W.D. & Hampson, S.E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19,* 242-279.

Milton, F.N. (2006). *Category construction: A study of the principles underlying category formation.* Unpublished doctoral dissertation, University of Exeter.

Milton, F.N., & Wills, A.J. (2004). The influence of stimulus properties on category construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 407-415.

Milton, F.N., Longmore, C.A., & Wills, A.J. (in press). Processes of overall similarity sorting in free classification. Journal of Experimental Psychology: Human Perception and Performance.

Overall, J.E. (1980). Power of chi-square tests for 2 x 2 contingency tables with small expected frequencies. *Psychological Bulletin, 87,* 132-135.

Pothos, E.M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science, 26*, 303-343.

Regehr, G. & Brooks, L.R. (1995). Category organization in free classification: The organizing effect of an array of stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 347-363.

Rosch, E. & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-605.

Smith, L.B. & Kemler, D.G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology, 10*, 502-532.

Spalding, T.L & Murphy, G.L. (1996). Effects of background knowledge on category construction. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 22, 525-538.

Ward, T.B. (1983). Response tempo and separable-integral responding: Evidence for an integral-to-separable processing sequence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 103-112.

Wills, A.J., & McLaren, I.P.L. (1998). Perceptual learning and free classification. *Quarterly Journal of Experimental Psychology, 51B*, 235-270.

Wittgenstein, L. (1958). *Philosophical Investigations* (2nd ed.). Oxford, England: Blackwell.