類別變異效果
The Category Variability Effect in Category Learning

吳 岳勳[†]，楊 立行[†]，王 曉薇[‡]
Yueh-Hsun Wu, Lee-Xieng Yang, Hsiao-Wei Wang

[†]政治大学，[‡]東北大學
National Chengchi University, Tohoku University
lxyang@nccu.edu.tw

Abstract

The category variability effect is referred to that when the two contrasting categories are not equally varying along the stimulus dimension, the very middle item in between the categories would be more likely classified to the category of high variety category although it is equally similar to the edge items of the categories. This phenomenon was firstly found with semantic stimuli, however, the following-up studies, with the perceptual stimuli, did not get this finding reliably. The present study is aimed to examine under which circumstances the category variability effect would occur. Instead of the visual stimuli, the experiments in this study use auditory stimuli (i.e., tones), for the auditory information (1-D information) would produce less noise in perception. Experiment 1 examined whether the category variability effect would be induced by a hint about the difference on category variability. The results showed no effect of the hint and also no category variability effect. Experiment 2 suggested the failure in Experiment 1 might result from an improper manipulation of category variability. In Experiment 2, we shrank the range of the low-variability category and found strong category variability effect when the low-variability category variability became even small. In order to examine the reliability of this observation, in Experiment 3, we turned to change the variability of the high-variability category. The results confirmed a strong category variability effect. Thus, it is suggested that this effect can also be found in the perceptual category learning task. Further, we developed an index *D* to measure the ratio between the category deviations. When *D* smaller than 1, the category variability effect is more likely to occur and disappear otherwise. With *D*, we can predict the occurrence of category variability effect.

Keywords — Category Variability Effect, Category Learning

Introduction

Normally, categorization is regarded as a process to put together similar objects in a category and separate dissimilar objects into different categories. In the past decades, a lot of theories posit that similarity is the basis for categorization (Kruschke,1992; Love, Medin, & Gureckis, 2004; Medin & Schaffer, 1978; Nosofsky,1986,1987). Although there are many different ways to interpret similarity, in this manuscript, we adopt Nosofsky’s (1986, 1987) definition for similarity as a negative exponential function of the distance between objects in a psychological space. That is, the closer two objects in a psychological space, the more similar they are to each other. Therefore, an object would be more likely to be classified to the category whose exemplars are relatively similar to it. This assumption is shared among all exemplar-based models, such as ALCOVE (Kruschke, 1992) and GCM (Nosofsky1986, 1987).

However, similarity might not be the sole basis for categorization, as revealed in the seminal study of (Rips, 1989). In his seminal study, the participants were asked to imagine a 3-in circular object and judge whether it is more similar to a quarter or a pizza as well as whether it is belonged to the category of coin or the category of pizza. According to the participants' estimation of the diameter lengths of different coins and pizzas in the pilot study, 3 inch was right in the middle between the edge exemplars of the two categories and most coin members had a diameter closer to 3 inches than the pizza members did. The results showed that the participants tended to judge this critical item[1] as more similar to coin, however,

---

[1] From now on, an object equidistant to the edge

they tended to classify it to the opposite category - pizza.

The most appealing part of this finding is why the critical item is not classified to the category whose members are obviously more similar to it, as suggested by the exemplar-based models. Among those psychologists who tried to provide a plausible explanation, Smith and Sloman (1994) noticed that the coin category is relatively less varying in diameter than the pizza category is. Thus, they suggested that the participants in Rips study might have applied some sort of rule in this task, such as "anything larger than 1 inch in diameter is not a coin". Accordingly, this finding has been regarded as evidence against the exemplar-based models. Given that this phenomenon seems to be quite relevant to the differential variability of two categories, it is fair enough referred to the category variability effect.

Although the category variability effect was found consistently with the semantic materials (e.g., the written sentence describing coin and pizza) (Rips & Collins, 1993; Smith & Sloman, 1994), it seemed to be very unreliable in the perceptual category learning task. For instance, Cohen, Nosofsky, and Zaki (2001) tried to replicate the Rips finding with straight lines of different lengths. These lines were defined as two categories by their line length. These authors asked the participants to learn the category membership in a trial-and-error fashion. The critical item was presented to the participants in the transfer phase. Although the results showed that the critical item was more likely to be classified to the high-variability category when the variability of that category increased, the mean probability to be as high-variability category was about .47 (not too much different from .50). Also, the low-variability category these authors used consisted of one exemplar only, that makes their task more like an identification task otherwise. Thus, the evidence for the category variability effect in a perceptual category learning task is still weak.

exemplars of two categories is called the critical item for the sake of writing convenience.

Similarly, Stewart and Chater (2002) tried to replicate Rips' (1989) experiment with the stimulus consisting of a circle with a dot attached to it. In their study, the high- and low-variability category respectively corresponded to different ranges over the dot position: one is larger than the other. Although these authors did observe the category variability effect, this result only occurred when the participants noticed the difference between two categories on the category variability in the way of seeing all stimuli of two categories at once while learning them. However, showing all stimuli at once to participants is not a conventional way in the category learning paradigm. Again, the category variability effect is still uncertain in the perceptual category learning task.

Hsu and Griffiths (2010) asked different groups of participants to learn categories of perceptual stimuli in two conditions. In the discriminative condition, the participants were given a stimulus and asked to judge to which category it was belonged. After the response was made, a corrective feedback was provided. In the generative condition, the participants were given the stimulus together with the category label to learn on each trial. The category structure and the experimental procedure were all the same in these two conditions. The results showed that the category variability effect occurred in the generative condition not in the discriminative condition. This result implies when the category characteristic (i.e., how varying the category is) is considered on classifying a stimulus, the category variability effect occurs.

From the previous review, it is clear that (1) the category variability effect can occur in the perceptual category learning task but that (2) it needs some particular experimental regimes (e.g., showing all stimuli at once or adopting observational learning alike paradigm) to induce. Why the feedback-learning paradigm cannot induce this effect becomes the major concern of this study. Perhaps, the ease of discriminating different stimuli perceptually is worth considering. In the previous reviewed studies, most studies except the work of Cohen et.al. (2001) did not provide the information about how close the stimuli were to each other in the psychological space. It is quite possible that the category variability effect does not occur because people cannot perceptually tell the difference between the stimuli of two categories (e.g., all dots of the circular stimuli in Stewart and Chater (2002) experiments only vary in a small region). Also, the stimulus material may share a part of discrimination ease. As the visual stimuli provide 2-dimensional information, which may relatively easily help the use of the unexpected strategy in learning categories. Thus, in order to examine whether the category variability effect can occur in the perceptual category learning task, (1) we adopt auditory material (e.g., tones) as stimulus in this study and (2) we systematically manipulate the psychological distance between the categories. Our hypothesis is that the category variability effect is a fundamental phenomenon, which does not need any top-down strategy to induce, as long as the perceived difference between the category variability is large enough.

Experiment 1

In this experiment and the rest experiments of this study, we used tones of different frequencies as stimuli. All tones were transferred from physical frequency f to psychological unit mel - mel scale by $m = 1127\log_e\left(\frac{f}{700}+1\right)$ (Steinberg, 1937; Stevens, Volkmann, & Newman, 1937). The mel scale is an interval scale which guarantees a good control over the psychological intensity of the tone stimulus we provided to participants. In Experiment 1, we put our emphasis on whether the top-down instruction is necessary for the occurrence of the category variability effect. The tones in the low-variability category varied from 480 mel to 520 mel with 10 mel as the interval. The tones in the high-variability category varied from 670 mel to 970 mel with 75 mel as the interval. The experiment contained two conditions depending on whether the participants were given a hint or not about the difference on the category variability.

Participants and Apparatus

All participants in this experiment were the undergraduate students in National Chengchi University (n=39 in the hint condition and n=38 in the no-hint condition). After testing, each participant would be reimbursed ($100) for effort and traffic expense. The experiment was conducted in a quiet booth. The experimental procedure, the displaying of stimuli, and the recording of response time and accuracy were all controlled by a MATLAB program implemented by the Psychophysics toolbox (Brainard, 1997; Pelli, 1997).

Stimuli and Procedure

All stimuli can be grouped to two categories with respect to the variety. In Figure 1, 10 learning stimuli (5 in each category) were shown in white bars and 7 testing stimuli were shown in black bars. In each condition, there were 5 leaning blocks each followed by 1 transfer block. In no matter learning or transfer block, each stimulus was presented twice randomly. The stimuli were the same for these two conditions, however, in the hint condition, the participants would be instructed that one category varied more than the other whereas the participants in the no-hint condition were just asked to learn the categories by trial-and-error. In the transfer blocks, the participants were asked to classify the transfer stimuli without any feedbacks.

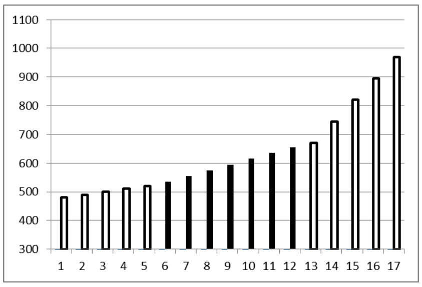


Figure 1. The category structure used in Experiment 1. The X axis denotes the number of the stimulus and the Y axis denotes the intensity of each stimulus tone in mel.

Results

The overall averaged accuracy of all participants was .89 which showed that the participants did learn the categories. The probability of the transfer item to be classified to the high-variability category was the main dependent variable. Hence, we averaged the response onto the transfer items in the last 3 blocks and conducted an Item (7) x Condition (2) mixed design ANOVA. The results showed a significant main effect of Item, $F(6,450) = 177.52$, $MSe = 0.02$, $p < .01$, no significant main effect of Condition, $F(1,75) < 1$, and no interaction effect, $F(6,450) = 1.83$, $MSe = 0.02$, $p = .09$. The no. 9 stimulus (595 mel) was the critical item as it was equidistant to the edge exemplars of the two categories. The probability of high-variability category on it was no different between the two conditions ($t(75) = -0.56$, $p = .26$). Also, the no. 9 stimulus was not significantly different from .50, ($t(38) = -1.23$, $p = .23$ for the hint condition and $t(37) = -0.318$, $p = 0.75$ for the no-hint condition). The probability of high-variability category on each transfer item can be seen in Table 1 and the standard error was shown in the parentheses. To sum up, Experiment 1 did not find any influence of hint on categorization performance, nor the category variability effect with or without hint.

Table 1. The probability of high-variety category on the transfer item

	535mel	555mel	575mel	595mel	615mel	635mel	655mel
Hint	.11(.02)	.17(.03)	.33(.04)	.46(.04)	.55(.04)	.67(.04)	.78(.03)
No-Hint	.19(.03)	.23(.04)	.39(.04)	.49(.04)	.56(.04)	.67(.03)	.73(.04)

Experiment 2

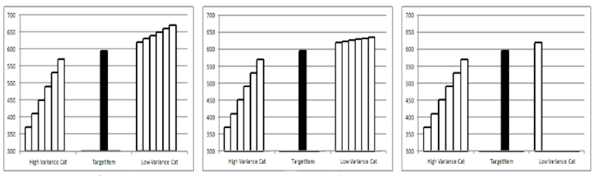
Experiment 1 showed no support for the category variability effect. This result is consistent with the report of Hsu and Griffith (2010) that the category variability effect did not occur in the discriminative condition (i.e., the feedback learning condition). However, this failure to observe the category variability effect might as well result from the improper manipulation of stimuli. In this experiment, we manipulated the range of the low-variability category as 50 mel, 15 mel, and 0 mel with the range of the high-variability category remained the same. Same as in Experiment 1, this experiment adopted the feedback learning paradigm. If the feedback learning condition cannot induce the occurrence of the category variability effect, then the range of low-variability category should not influence the probability of the critical item to be as the high-variability category, otherwise it would.

Participants and Apparatus

There were 42, 37, and 37 undergraduate students in National Chengchi University recruited respectively in the conditions of LR50, LR15, and LR0. Same as in Experiment 1, the experiment was conducted on the IBM compatible PC under the control of a MATLAB program.

Stimuli and Procedure

The stimuli used in all conditions can be seen in Figure 2. The three conditions were shown from left to right. In each condition, there were 6 learning stimuli in each category. The high-variability category always ranged from 370 mel to 570 mel in all conditions, whereas the low-variability category has different ranges in different conditions (620 – 670 mel in LR50, 620 – 635 mel in LR15, and 620 – 620 mel in LR0). The participants were asked to learn the categories in 5 learning blocks, each of which was followed by a transfer block. In each learning block, the 12 learning stimuli were presented twice (12 x 2 = 24 trials) in a random order. In each transfer block, in addition to the critical item (590 mel), 4 learning stimuli were also presented, which were randomly selected from both categories (2 from each category).



LR50 LR15 LR0

Figure 2. The category structure in Experiment 2.

Results

The probability of high-variability category on the critical item increased as the range of the low-variability category decreased, .56 (LR50), .86 (LR15), and .89 (LR0). The one-way between subject ANOVA showed a significant effect of Condition on the critical item [$F(2,113) = 19.99$, $MSe = 0.07$, $p < .01$]. The simple comparison test supported that the critical item was less likely classified to the high-variability category in LR50 than LR15 [$F(2,113) = 26.57$, $MSe = 0.07$, $p < .01$] and LR0 [$F(2,113) = 31.56$, $MSe = 0.07$, $p < .01$]. With no surprise, the tendency to classify the critical item to the high-variability category in LR50 was no different from a chance level [$t(41) = 1.22$, $p = .23$], whereas it was much larger than .50 in LR15 [$t(36) = 9.36$, $p < .01$] and LR0 [$t(36) = 11.752$, $p < .01$]. Thus, we observed the category variability effect in this experiment.

Although Cohen et.al. (2001) observed an increased probability of high-variability category on the critical item by making the high-variability category more varying, not like us by making the low-variability category less varying, both findings converge on a suggestion that the category variability effect occurs only when the difference on the category variability between two categories is large.

More importantly, the success of this experiment clearly indicates that the category variability effect can be observed with the feedback learning paradigm.

Experiment 3

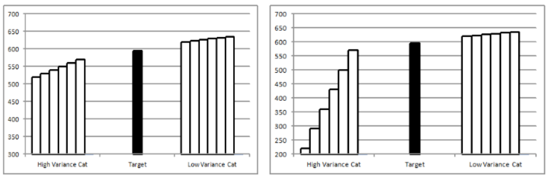
Although the category variability effect is strong in Experiment 2, it is still worth checking if this effect is independent of the mel intensity and only relevant to the category structure. Thus, in this experiment, the variability of the stimuli of low mels, namely the high-variability category in Figure 2, would be manipulated. While keeping the low-variability category always in the range of 15 mels (LR15), the range of the high-variability category in the two conditions were 350 (HR350) and 50 (HR50) respectively. The rest procedure and experiment design were all the same as in the previous experiments.

Participants and Apparatus

There were 39 undergraduate students in National Chengchi University participating in the HR50 condition and 38 in the HR350 condition. They would be reimbursed with $50 after testing for their effort and traffic expense. Same as in the previous experiments, the experiment was conducted on the IBM compatible PC in a quiet booth. The displaying of stimulus, the recording of response time and accuracy, and the whole experimental process were all under control of a MATLAB program.

Stimuli and Procedure

The stimuli can be seen in Figure 3. The left panel shows the category structure in the HR50 condition, in which the range of the high-variability category is 50 mels from 520 to 570 mels. The right panel shows the category structure in the HR350 condition, in which the range of the high-variability category is 350 mels from 220 to 570 mels. The low-variability category in either condition has a range of 15 mels from 620 to 635 mels. There were 6 learning stimuli in each category and 1 critical item (590 mel) in between the two categories.



HR50 HR350

Figure 3. Category structure used in Experiment 3.

Same as the procedure of the previous experiments, the participants were asked to learn stimuli in 5 blocks, each of which was followed by a transfer block. Therefor, there were 5 transfer blocks as well. In each learning block, 12 stimuli from both categories were presented twice each in random. In the transfer block, in addition to the critical item, 4 learning stimuli randomly selected from the categories were presented to the participants. In the learning blocks, the participants would be given a corrective feedback as “correct” or “wrong” after a response was made. However, in the transfer blocks, the participants would not get any feedback.

Results

In total, there were 10 blocks (learning plus transfer blocks) in this task and the averaged accuracy for all participants were .95. This shows that the participants can learn this category structure very well. However, there were three participants who could not get a learning accuracy as high as .80. These participants’ data thus were not included in the further data analysis. In total, 39 data sets in the HR50 condition and 35 data sets in the HR359 condition were analyzed.

The mean probability of high-variability category on the critical item was .87 in the HR50 condition and .78 in the HR350 condition. Both mean accuracies were significantly higher than the 50% chance level, [$t(38) = 11.38$, $p < .01$ for the HR50 condition; $t(34) = 6.742$, $p < .01$ for the HR350 condition].

In both conditions, the critical item was significantly classified to the high-variability category, thus the category variability effect was confirmed again. Therefore, it is supported that the mel intensity has nothing to do with the category variability effect.

For a more thorough understanding of the category variability effect, we put together in Table 2 the participants’ performance on the critical item from the conditions in Experiment 2 and 3 in this study, as these conditions were conducted in the same regime.

Table 2. The probability of high-variability category on the critical item in the conditions of Experiment 2 and Experiment 3. The standard error was shown in the parentheses.

LR50	LR0	LR15	HR50	HR350
0.56(0.08)	0.88(0.04)	0.86(0.04)	0.87(0.03)	0.78(0.04)

Apparently, except in the LR50 condition, the category variability effect occurred in all other conditions. This inspection is supported by the analysis of one-way between-subject ANOVA, $F(4,188) = 12.90$, $MSe = 0.08$, $p < .01$. The Scheffe test shows the probability of high-variability category on the critical item in LR50 is significantly less than that in the other conditions, (LR50 vs. LR0: $F(4,185)$

= 35.24, MSe = 0.06, p < .01; LR50 vs. LR15: F(4,185) = 29.67, MSe = 0.06, p < .01; LR50 vs. HR50: F(4,185) = 32.87, MSe = 0.06, p < .01; LR50 vs. HR350: F(4,185) = 14.94, MSe = 0.06, p < .01). The other four conditions are not significantly different from one another (LR0 vs. LR15: F(4,185) = 0.23, MSe = 0.06, p = .99; LR0 vs. HR50: F(4,185) = 7.69, MSe = 0.06, p > .99; LR0 vs. HR350: F(4,185) = 3.7, MSe = 0.06, p > .99; LR15 vs. HR50: F(4,185) = 4.13, MSe = 0.06, p > .99; LR0 vs. HR350: F(4,185) = 2.12, MSe = 0.06, p = .71; HR15 vs. HR350: F(4,185) = 2.81, MSe = 0.06, p = .59).

Measurement of Category Variability Effect

It is clear in Table 2 that the category variability effect does not occur only in the LR50 condition and does in the rests. To tell the difference between these conditions quantitatively, we generated an index called D index, $D = I_L SD_L / SD_H$, where I_L is the interval of the low-variability category and SD is the standard deviation of the category. The basic idea of this equation is that the more the two categories differ in variability, the more likely the category variability effect would occur. The interval of the low-variability category is a scaling constant to make D = 1 as a boundary for the occurrence of category variability effect. Thus, the category variability effect occurs when $D < 1$, and disappears when $D > 1$ otherwise.

Therefore, we apply the D index as the measurement of the category variability category for our experiments as well as the experiments we review. For the reason that we want to get a fair comparison between these data in D, we do not include the experiments not using feedback learning or using some extra instruction in the attempt to induce the category variability effect.

The results are shown in Table 3. The probability of high-variability category on the critical item, p, is the behavioral index for the extent of category variability effect, the larger the stronger the category variability effect is.

Table 3. The D index of different experiments

Experiment	SD_L	SD_H	I_L	D	p
No-Hint	14.14	106.07	10	1.33	.49
LR0*	0.00	56.57	0	0.00	.88
LR15*	4.24	56.57	3	0.22	.86
LR50	14.14	56.57	10	2.50	.56
HR50*	4.24	17.07	3	0.75	.87
HR350*	4.24	119.55	3	0.11	.78
Cohen(2001)	0.00	25.62	0	0.00	.47
Stewart(2002)	11.00	28.00	5	1.96	.39
Sakamoto(2006)*	3.42	34.16	2	0.20	.69
Hsu(2009)	14.14	106.07	10	1.33	.47

Basically the D index runs well for almost all experiments we review, denoted by * in the column of experiment in Table 3. As expected, when $D > 1$, p is not too large. That means a weak tendency for the category variability effect. On the contrary, when $D < 1$, p gets larger, a relatively strong tendency for the category variability effect. However, this regularity is not that stable for the case in which there is only one item in the low-variability category ($I_L = 0$, hence $D = 0$). In table 3, for the case of LR0 in our experiment there is a strong category variability effect, $p = .88$, whereas for the case of Cohen et.al. (2010) experiment, there is a weak category variability effect, $p = .47$. Probably this is because that the case with only one item in a category makes the whole learning task like an identification task rather than a category learning task and some unexpected strategy might be involved in making responses as well.

To sum up, the category variability effect does occur in the feedback learning task with the perceptual stimuli. Our results suggest that the inconsistent findings in the past studies might come from the improper manipulation of the stimuli. Further, we provide a quantitative index *D* to measure the possible extent of the category variability effect given the category structure.

General Discussions

In this study, the main concern is to examine whether the category variability effect would occur in the perceptual category learning task. For this aim, unlike the precedent research, we used the auditory tones as stimuli in the attempt to get a better control on the psychological intensity of the stimuli. Three experiments were conducted in order to pursue this issue.

In Experiment 1, the participants learned the categories in a feedback learning paradigm with a hint or no hint about the categories varied differently. The results showed no category variability effect on the critical item, no matter the hint was provided or not. However, this failure might result from the improper manipulation of the stimuli. In Experiment 2, we shrank the range of the low-variability category to make three conditions (LR50, LR15, and LR0), which covered the higher region on the mel scale. The results showed a clear category variability effect in the conditions of LR15 and LR0, but not in the condition of LR50. This experiment suggested that in a traditional learning paradigm (i.e., feedback learning), the category variability effect could occur. In order to examine the reliability of this observation, in Experiment 3, we instead manipulated the range of high-variability category, which covered the lower region on the mel scale. The category variability effect was observed still. Thus, these results provide a positive support for our hypothesis that the category variability effect can occur in the perceptual category learning task.

Further, in order to get a comprehensive understanding about why the category variability effect occurs in some circumstance but not in others, we developed a quantitative index *D*. *D* actually represents the extent of ratio between category deviations. As shown in Table 3, *D* provides a good account for the experiments on this issue.

Strategy use and category variability effect

The study of Hsu and Griffith (2010) implies that the occurrence of category variability effect results from the use of a top-down strategy, in that people infer to which category the current stimulus is belonged according to their knowledge about the categories. However, our study suggests that the stimulus-response associations gained in the feedback learning paradigm is sufficient to induce the category variability effect. This is not to say that we exclude the possibility of the use of top-down strategy. What we suggest is just that the occurrence of category variability effect has a more perceptual account.

Category variability and category learning model

Although many studies have shown that the exemplar-based model has difficulty accounting for the category variability effect Cohen Johansen, the other models, which are evident to be able to account for this effect seem to be post hoc. For instance, the model of Sakamoto, Love, and Jones (2006) simply adds a component for category SD to account for the category variability. Some others like the GRT model (Ashby & Gott, 1988), although not that post hoc, still cannot provide an explanation to how the information of category variability is transferred so as to get the category variability effect. Therefore, it is worth pursuing this issue, specifically with the confidence for the occurrence of this effect in advance. Our D index provides a clear prediction to the occurrence of category variability effect that in turn can help other researchers continue studying this issue with a reliable measurement.

References

Ashby, F. G., & Gott, R. E. (1988). Decision rule in the perception and categorization on multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 14, 33-53.

Brainard, D. H. (1997). The psychophysics toolbox. Spatial Vision, 10, 433-436.

Cohen, A., Nosofsky, R., & Zaki, S. (2001). Category variability, exemplar similarity, and perceptual classification. *Memory & Cognition, 29*(8), 1165-1175.

Hsu, A. S., & Griffiths, T. L. (2010). Effects of generative and discriminative learning on use of category variability. *Proceedings of 32nd Annual Conference of the Cognitive Science Society.*

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

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of categorization. *Psychological View,* 111, 309-332.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*(3), 207-238.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39-57.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and classification of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 13, 87-108.

Pelli, D. C. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision,* 10, 437-442.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning.* (pp. 21-59): New York, NY, US: Cambridge University Press.

Rips, L. J., & Collins, A. (1993). Categories and resemblance. *Journal of Experimental Psychology: General, 122*(4), 468-486.

Sakamoto, Y., Love, B. C., & Jones, M. (2006). Tracking Variability in Learning: Contrasting Statistical and Similarity-Based Accounts. *Proceedings of Cognitive Science Society*, Mahwah, NJ.

Smith, E. E., & Sloman, S. (1994). Similarity- versus rule-based categorization. *Memory & Cognition, 22*(4), 377-386.

Steinberg, J. C. (1937). Positions of Stimulation in the Cochlea by Pure Tones. *The Journal of the Acoustical Society of America, 8*(3), 176-180.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America, 8*, 185-190.

Stewart, N., & Chater, N. (2002). The effect of category variability in perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(5), 893-907.