Title

Less is More: Stimulus Feedback Co-Occurence in Perceptual Category Learning

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN

1069-7977

Authors

Maddox, W. Todd
Markman, Arthur
Worthy, Darrell A.

Publication Date

2009

Peer reviewed

Less is More: Stimulus-Feedback Co-Occurrence in Perceptual Category Learning

Darrell A. Worthy (worthyda@mail.utexas.edu)
Department of Psychology, 1 University Station A8000
Austin, TX 78712-0187 USA

W. Todd Maddox (maddox@psy.utexas.edu)
Department of Psychology, 1 University Station A8000
Austin, TX 78712-0187 USA

Arthur B. Markman (markman@psy.utexas.edu)
Department of Psychology, 1 University Station A8000
Austin, TX 78712-0187 USA

Abstract

This study examined the effects of stimulus-feedback co-occurrence on rule-based and information-integration category learning. Rule-based categories are those for which a verbalizable rule is optimal. Information-integration categories are those for which the optimal rule is non-verbalizable. Participants performed a rule-based or an information-integration task where the stimulus co-occurred with the feedback (Stimulus Present) or was removed prior to feedback presentation (Stimulus Absent). Previous research examining the neural substrates of rule-based and information-integration category learning suggests that stimulus-feedback co-occurrence should support rule-based learning, but should harm information-integration learning because it will increase the prevalence of rule use. This prediction was confirmed in the current study. Implications for theories of category learning are discussed.

Keywords: Category learning; Feedback; Procedural Learning

Introduction

There is nearly unanimous agreement that category learning is mediated by multiple, distinct neural systems (e.g. Ashby, Alfonso-Reese, Turken, & Waldron, 1998). Even so, much is still unknown about the processing characteristics of each system. Two distinct neural systems have been the focus of much research and are thought to mediate different types of category-learning. An explicit *hypothesis-testing* system is thought to mediate rule-based category learning, and relies on frontal brain regions and the head of the caudate nucleus. An implicit *procedural-learning* system is thought to mediate information-integration category learning, and relies on the body and tail of the caudate nucleus. Rule-based learning involves testing verbalizable rules in order find the rule that optimally separates stimuli into the categories. Information-integration learning involves associating regions of perceptual space with actions that lead to reward (e.g. Maddox, Ing, & Bohil, 2004; Ashby, Ell, & Waldron,, 2003; Spiering and Ashby, 2008) The hypothesis testing system uses executive resources to construct and test verbalizable rules, whereas the procedural system uses dopamine-mediated reward learning to associate regions of the stimulus space with a response.

Figure 1 shows stimuli from rule-based and information-integration category structures. The stimuli are sine-wave gratings (Gabor patches) that vary in their spatial frequency and spatial orientation. Figure 1a shows a rule-based category structure. Here the optimal rule is to classify stimuli with relatively low spatial frequency into one category and stimuli with relatively high spatial frequency into the other category. In contrast, there is no verbalizable rule that can be employed to optimally distinguish the categories in the information-integration structure shown in Figure 1b. Here the optimal rule may be verbalized as: 'if the spatial orientation is greater than the spatial frequency, classify the stimuli into one category, but if the spatial orientation is less than the spatial frequency, classify the stimuli into the other category.' However, this type of rule is difficult or impossible to implement to solve the task because the relative magnitude of each spatial dimension cannot be easily compared.

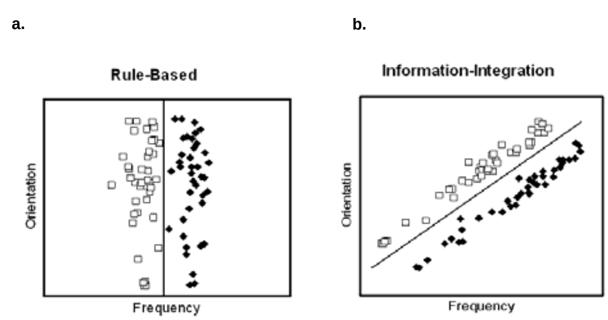


Figure 1: (a) Plot of stimuli from the rule-based category structure. (b) Plot of stimuli from the information-integration category structure.

Previous research suggests that the hypothesis testing system relies on executive attention and is not vulnerable to manipulations of feedback delay (e.g. Zeithamova &

Maddox, 2006; Maddox Ashby, & Bohil, 2003 Maddox & Ing, 2005). In contrast, the procedural system does not require executive attention, but requires a short feedback delay to optimize learning.

The COmpetition between Verbal and Implicit Systems (COVIS, Ashby et al., 1998) model proposes that the hypothesis-testing system (i.e. verbal) competes with the procedural (i.e. implicit) system to determine which of these two systems governs responding on a given trial. Although both systems are thought to be active on each trial, there is an initial bias toward the hypothesis testing system. Control is gradually passed to the procedural system if the responses it generates become more accurate. If the hypothesis testing system is highly accurate then it may take longer for control to be passed to the procedural system. One possibility (suggested in COVIS) is that the hypothesis-testing system acts as a type of gating mechanism for the procedural system. When the hypothesis-testing system is performing well, it governs responding. When the hypothesis-testing is performing poorly, control is passed to the procedural system.

A general hypothesis that follows from COVIS is that conditions that enhance the hypothesis-testing system also harm the procedural system. Maddox and colleagues recently showed that feedback properties that led to more accurate rule-based classification led to less accurate information-integration classification (Maddox, Love, Glass, & Filoteo, 2008). They proposed that full feedback, as opposed to partial feedback, would lead to better performance on rule-based tasks, but worse performance on information integration tasks. Participants performed either a rule-based or an information-integration task with either full or partial feedback. In this experiment the stimuli were from one of four categories. In the full feedback condition participants were told whether they were correct, and also what category the stimulus belonged to. In the partial feedback condition participants were told whether they were correct, but were not told which category the stimulus belonged to. They found that more information given during feedback led to better performance on rule-based tasks, but worse performance on information-integration tasks.

Another factor that could increase the reliance on the hypothesis-testing system is stimulus-feedback co-occurrence. The hypothesis-testing system operates by testing verbalizable rules that can distinguish stimuli from each category. When feedback is given the classifier must determine why the use of a given rule did or did not lead to a correct categorization of the stimulus. If the stimulus is present during feedback then the hypothesis testing system can continue to test rules during feedback. However, if the stimulus is removed prior to feedback then it should make rule-use more difficult. The classifier must hold an image of the stimulus, as well as the verbalizable rule used during that trial in working memory in order to process the feedback. Stimulus-feedback co-occurrence should make the hypothesis-testing system easier to use and more efficient.

An interesting prediction from this theory is that stimulus-feedback co-occurrence may lead to an increased reliance on the hypothesis-testing system. This would come at the expense of the procedural system. However, when the stimulus and the feedback do not co-occur the hypothesis-testing system may be abandoned sooner in favor of the procedural system. Stimulus-feedback co-occurrence may lead to better performance on rule-based tasks, but may actually harm performance on information-integration tasks.

In this paper we test the hypothesis that stimulus-feedback co-occurrence improves performance on rule-based tasks, but hinders performance on information-integration tasks. If the stimulus is present on the screen for the duration of the feedback presentation then this should enhance the hypothesis-testing system. Having the stimulus present during feedback allows the participant to further examine the stimulus to determine why the rule they used to classify the stimulus did or did not work. This should also reduce the working memory demands of the hypothesis-testing system because the properties of the stimulus and the rule-used on that trial do not have to be recalled during feedback. However, when the stimulus is not present during feedback control may more rapidly shift from the hypothesis-testing system to the procedural system because the hypothesis testing system may have to work harder to recall the stimulus properties and the rule-used while processing the feedback. This may reduce the amount of explicit rule use, and thus hurt performance on rule-based tasks but help performance on information-integration tasks.

Experiment

Method

Eighty participants from the University of Texas community were given course credit or monetary compensation to participate in the experiment. Participants were randomly assigned to one of four between-subjects conditions that consisted of the factorial combination of two category types (Rule-Based vs. Information-Integration) and two types of feedback presentation (Stimulus Present vs. Stimulus Absent).

Stimuli

The stimuli for the two category structures are plotted in Figure 1. Stimuli were Gabor patches that varied in the frequency of the bars and their orientation relative to the computer screen.

Procedure

Participants performed five blocks of 80 trials. On each trial a stimulus was presented on the screen and participants pressed a key to indicate which category they thought the stimulus belonged to. Participants were given as long as they wished to make a response. For the Stimulus Absent condition the stimulus was removed from the screen immediately after a response was made. Feedback was given 500ms after the response was made. If the choice was correct then the word "Correct" appeared at the bottom of the screen. If the choice was incorrect then the phrase "No, that was a 1 (or 2)" appeared at the bottom of the screen. Feedback was shown for 3500ms. The screen then went blank and the next trial began. The procedure was identical for the Stimulus Present condition except that the stimulus did not disappear after the response was made. Instead, the stimulus stayed on the screen for the duration of the feedback presentation and was only removed upon the beginning of the next trial.

Results

Performance Measures

Figure 2 shows the mean accuracy for each condition averaged across all blocks. The data were subjected to a 2 (Category Type) X 2 (Feedback Type) X 5 (Block) repeated measures ANOVA. There was a significant effect of block $F(4)=43.40$, $p<.001$, $\eta^2=.36$. There was also a significant Category Type X Feedback Type interaction, $F(1,76)=11.47$, $p<.01$, $\eta^2=.13$. To examine the nature of the interaction we compared the performance of participants across the four conditions. Participants in the Rule-Based Stimulus Present condition (M=.79) performed marginally better than participants in the Rule-Based Stimulus Absent Condition (M=.71), $F(1,38)=3.32$, $p<.10$, $\eta^2=.08$. In contrast, participants in the Information-Integration Stimulus Absent condition (M=.74) performed significantly better than participants in the Information-Integration Stimulus Present condition (M=.64), $F(1,38)=10.30$, $p<.01$, $\eta^2=.21$.

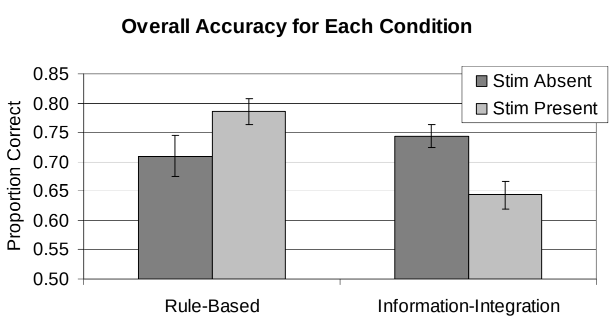


Figure 2: Proportion correct for each condition averaged across all blocks

Model-Based Analyses

The accuracy-based analyses support the proposal that having the stimulus present during feedback leads to better performance on rule-based tasks but worse performance on information-integration tasks. Accuracy analyses are an informative measure of performance, but they provide no information regarding the specific strategies used to classify the stimuli into the categories. To address this issue we applied decision bound models (Maddox, 1999; Maddox and Ashby, 1993) separately to the data from each participant on a block by block basis. All of the analyses were performed at the individual-participant level because of concerns with modeling aggregate data (e.g., Estes, 1956; Maddox, 1999; Maddox & Estes, 2004; Smith & Minda, 1998).

Decision bound models assume that participants use a decision bound to separate stimuli into categories with stimuli on one side of the bound being classified into one category and stimuli on the other side of the bound being classified into the other category. The optimal decision bound in the Rule-Based condition is depicted by the vertical line in Figure 1a. Examples of sub-optimal strategies would be to shift the decision bound along the frequency dimension, or to place a decision bound along the orientation dimension. The optimal decision bound in the Information-Integration condition is depicted by the diagonal line in Figure 1b. Sub-optimal strategies might include altering the slope or y-intercept of the optimal decision bound, or setting a decision bound on either the frequency or orientation dimension.

Fits of the models will be useful in determining if stimulus presence during feedback leads to an optimal or suboptimal classification strategy. For example, the low accuracy rates for participants in the Information-Integration Stimulus Present condition may be due to a use of suboptimal rule-based strategies. Similarly, the lower accuracy rates for participants in the Stimulus Absent condition may be due to an inability to apply the optimal rule as a classification strategy.

Rule-Based Models

The *optimal unidimensional frequency model* assumes that the participant uses the optimal criterion along the spatial frequency dimension and applies the rule: "Respond 'A' if the spatial frequency is low and 'B' if it is high." This model has one free parameter that represents the variance of internal (perceptual and criterial) noise, and was fit only to data from participants performing rule-based tasks. The *generalized unidimensional frequency rule model* assumes that the participant uses a criterion along the spatial frequency dimension, but allows the criterion value to be estimated from the data (2 free parameters total). The *generalized spatial orientation model* assumes that the participant uses a criterion along the spatial orientation dimension, and allows the criterion value to be estimated from the data (2 free parameters total). These two models were fit to data from all participants.

The *conjunction models* assume that the participant uses a conjunctive rule in which he or she makes separate decisions about the levels of the two dimensions and then selects a response based on the outcome of these two decisions. Two conjunctive rules were examined:

1. "Respond 'A' if spatial frequency is low and orientation is large, otherwise respond 'B,'" and
2. "Respond 'B' if spatial frequency is high and orientation is small, otherwise respond 'A'."

Both rules partition the perceptual space into four regions. The first assigns one to Category A and three to Category B, and the second assigns three to Category A and one to Category B. The conjunction models have three parameters (a criterion on each dimension, and an internal noise parameter). The conjunction models were fit to data from all participants.

Information-Integration Models

The *generalized linear classifier model* (GLC) assumes that the decision bound between each pair of categories is linear. This produces an information-integration decision strategy because it requires linear integration of perceived frequency and orientation. The GLC has three parameters: the slope and intercept of the linear bound, and an internal noise parameter. The *optimal GLC model* assumes that the participant uses the linear bound that maximizes accuracy. This model has only one free parameter representing the internal noise. The GLC was fit to all data, and the optimal GLC was fit only to data from Information-Integration participants.

Random Responder Model

Each participant's data was also fit by a one-parameter random responder model that assumed a fixed probability (estimated by the model) of responding "A" for all the stimuli.

Model Fitting Procedure

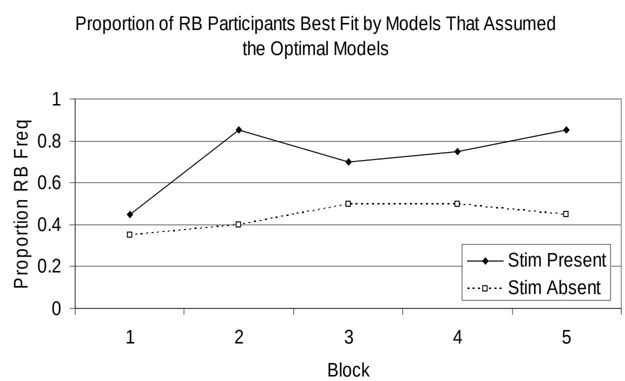
Each model was fit separately to the data from each of the eight blocks of trials in sessions four and five for each participant. The model parameters were estimated using maximum log-likelihood (Wickens, 1982), and the goodness-of-fit statistic used was $AIC = -2lnL + 2k$, where k is the number of free-parameters (Akaike, 1974). The AIC statistic penalizes models with extra free parameters. The best fitting model is the model with the smallest AIC value. For each block for each participant in each session we determined which model provided the best fit to the data.

Model Fitting Results

Within each category structure we compared the proportion of participants in the Stimulus Present and Stimulus Absent conditions who were fit by models of the same form as the optimal model based on the category structure. For participants in the Rule-Based condition this was either the optimal or generalized spatial frequency model. For participants in the Information-Integration condition this was either the generalized or optimal linear classifier model.

Figure 3 presents the proportion of participants in each condition who were fit best by the models that assumed use of the optimal strategy. For participants performing rule-based tasks a higher proportion of participants in the Stimulus Present condition were fit best by one of the optimal models in all five blocks (p=.03 by sign test, one tailed). For participants performing information integration tasks a higher proportion of participants in the Stimulus Absent condition were fit best by one of the optimal models in four of the five blocks.

a.



b.

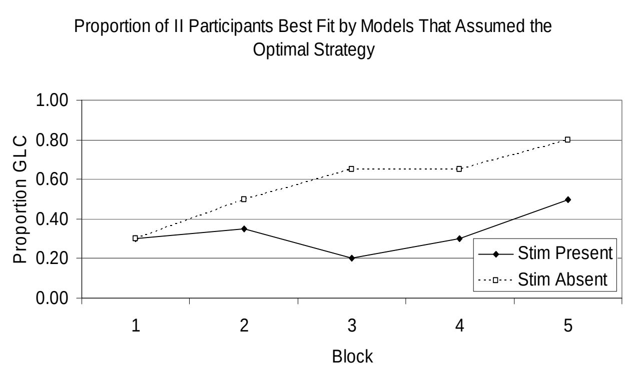


Figure 3: (a) Proportion of Rule-Based participants fit best by either the optimal or generalized spatial frequency model. (b) Proportion of Information-Integration participants fit best by either the optimal or generalized linear classifier model.

We performed binomial tests within each category structure for each block to compare the proportion of optimal model fits between the Stimulus Absent and

Stimulus Present conditions. For participants performing rule-based tasks a significantly higher proportion of data sets from participants in the Stimulus Present condition were fit best by the optimal models in block 2 (p<.01), block 4 (p<.05), and block 5 (p<.05) than data from participants in the Stimulus Absent condition. There was a marginally greater proportion of optimal model fits from data from participants in the Stimulus Present condition in block 3 (p<.10).

For participants performing information-integration tasks a significantly higher proportion of data sets from participants in the Stimulus Absent condition were fit best by one of the optimal models in block 3, block 4, and block 5 (all p<.01).

Discussion

Most theories of learning would predict that more information about the stimulus should lead to better learning. Participants in our Stimulus Present condition were allowed to continue processing the stimulus during the entire feedback interval. We proposed that this enhances rule-based learning because specific rules could be applied to the stimulus when full knowledge of the stimulus' category membership was available. However, this enhanced rule-based processing could delay the passing of control over categorization from the hypothesis-testing system to the procedural system.. As a result, performance for participants in the Information-Integration Stimulus Present condition should be worse than for those in the Stimulus Absent condition. Stimulus-feedback co-occurrence should reduce the amount of working memory needed for the hypothesis-testing system to operate. The participant does not have to hold stimulus information or the rule they implemented on that trial in working memory because the stimulus is present when feedback occurs. This should enhance the hypothesis-testing system and delay the passing of control from the hypothesis-testing system to the procedural system. When the stimulus and feedback do not co-occur the hypothesis-testing system should be at more of a disadvantage because stimulus information must be held in working memory. This should speed the passing of control from the hypothesis-testing system to the procedural system, which should enhance information-integration classification accuracy.

Our results offer strong support for our predictions. For participants performing rule-based tasks those in the Stimulus Present condition were more accurate and a greater proportion of their data sets were fit best by models that assumed the use of the optimal strategy (i.e., a rule-based model). However, for participants performing information-integration tasks those in the Stimulus Absent condition had higher accuracy rates and a greater proportion of their data sets were fit best by models that assumed use of the optimal strategy (i.e., an information-integration model).

The interaction between category type and stimulus presence during feedback offers more support for a multiple-systems view of perceptual category-learning. Specifically, there is abundant evidence for the two systems discussed here and first introduced by Ashby and colleagues (e.g. COVIS, Ashby et al., 1998). An important component of the COVIS model is the *competition* between the two systems. Both systems are thought to be active on each trial, but one system 'wins' the competition to govern responding.

An interesting corollary to this theory is that conditions which detract from one system's efficiency will enhance the other system's efficiency. Similarly, conditions which enhance one system's efficiency will detract from the other systems efficiency. In the current work we proposed that stimulus-feedback co-occurrence would enhance the hypothesis-testing system and, in turn, detract from the procedural system. The data strongly supported this prediction. We also predicted that the absence of a stimulus-feedback co-occurrence would detract from the hypothesis-testing system, and enhance the procedural system. This prediction was also strongly supported by the data.

Other work on the dissociation between rule-based and information-integration classification offers similar support for the trade-off between the two systems. Recent work in our labs has shown that certain motivational factors can enhance or detract from the hypothesis-testing system, and that the procedural system performs better or worse depending on the performance of the hypothesis-testing system. Maddox, Baldwin, and Markman (2006) showed that the alignment of short-term and long-term goal states (a 'regulatory fit') led to better rule-based accuracy, but worse information-integration accuracy. This was presumably because a regulatory fit led to an increase in executive resources which enhanced the hypothesis-testing system at the expense of the procedural system. Maddox et al. (2006) also showed that a misalignment of short-term and long-term goals (a 'regulatory mismatch') led to better information-integration accuracy, but worse rule-based accuracy (see also Grimm, Markman, Maddox, & Baldwin, 2008).

Markman, Maddox, and Worthy (2006) examined the effects of social pressure on rule-based and information-integration category-learning. They found the pressure led to worse rule-based accuracy, but better information-integration accuracy. The mechanism was again the same. Pressure detracted from the hypothesis-testing system which helped the procedural system. A lack of social pressure did not harm the hypothesis-testing system so control was not passed as quickly to the procedural system. This improved rule-based accuracy, but led to worse information-integration accuracy (see also Worthy, Markman, & Maddox, 2009).

Finally, Decaro, Thomas, and Beilock (2007) showed that individual differences in working memory influence performance on rule-based and information-integration category learning tasks. Individuals with high working memory capacity performed better on rule-based tasks than those with low working memory capacity. However, individuals with low working memory capacity performed better on information-integration tasks than those with high working memory capacity. This also supports the view that enhancement of one system comes at the expense of the other system. High-working memory capacity leads to an increased reliance on the hypothesis-testing system, which comes at the expense of the procedural system. The reverse is true for individuals with low working memory capacity.

This work offered strong support for a multiple systems view of category learning. It also points out that there is a tradeoff between the hypothesis-testing and procedural systems. More broadly, this work suggests a trade-off between explicit and implicit systems in a variety of situations besides categorization (e.g. Sloman, 1996). It is probable that these two systems are operative, and competing against one another, whenever humans are performing an action or making a decision. Future research on the trade-off between the two systems is important for fully understanding the complexities of human cognition and behavior.

Acknowledgments

This research was supported by NIMH grant MH077708 to WTM and ABM, and a supplement to NIMH grant MH077708 to DAW. We thank Bo Zhu, Maia Langford and all the research assistants in MaddoxLab for help in collecting the data. We also thank Matt Jones and Tyler Davis for valuable comments on the results.
Correspondence concerning this article should be addressed to W. Todd Maddox or Arthur Markman both at: University of Texas, 1 University Station A8000, Department of Psychology, Austin, Texas, 78712 (e-mail: maddox@psy.utexas.edu or markman@psy.utexas.edu.)

References

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442-481.

Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition, 31*, 1114-1125.

DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107*, 284-294.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, 53, 134-140.

Grimm, L.R., Markman, A.B., Maddox, W.T., & Baldwin, G.C. (2008). Differential effects of regulatory fit on category learning. *Journal of Experimental Social Psychology, 44*, 920-927.

Maddox, W. T. (1999). On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception and Psychophysics*, *61*, 354-374.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception and Psychophysics, 53*, 49-70.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003) Delayed feedback effects on rule-based and information-integration category learning. *Psychology and Aging. 19*, 171-182.

Maddox, W. T., Baldwin, G. C., & Markman, A. B. (2006). A test of the regulatory fit hypothesis in perceptual classification learning. *Memory & Cognition, 34*, 1377-1397.

Maddox, W. T., & Estes, W. K. (2004). Predicting true patterns of cognitive performance from noisy data. *Psychonomic Bulletin & Review*, *11*, 1129-1135.

Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural learning system but not the hypothesis-testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 100-107.

Maddox, W. T., Ing, A. D., & Bohil, C. J. (2004). Evidence for a procedural-learning-based system in perceptual category learning. *Psychonomic Bulletin & Review, 11*, 945-952.

Maddox, W. T., Love, B. C., Glass, B. D., & Filoteo, J. V. (2008). When more is less: Feedback effects in perceptual category learning. *Cognition, 108*, 578-589.

Markman, A. B., Maddox, & W. T., Worthy, D. A. (2006) Choking and excelling under pressure. *Psychological Science*. *17*, 944-948.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3-22.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*, 1411-1436.

Spiering, B. J., & Ashby, F. G. (2008) Response processes in information-integration category learning. *Neurobiology of Learning and Memory*, *90*, 330-338.

Worthy, D. A., Markman, A. B., & Maddox, W. T. (2009)**.** What is pressure? Evidence for social pressure as a type of regulatory focus. *Psychonomic Bulletin and Review*, *16*, 344-349.

Zeithamova, D., & Maddox, W. T. (2006). Dual task interference in perceptual category learning. *Memory & Cognition*, *34*, 387-398.