Title
Individual Variability in Strategies and Learning Outcomes in Auditory Category Learning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Roark, Casey L
Chandrasekaran, Bharath

Publication Date
2021

Peer reviewed

Individual Variability in Strategies and Learning Outcomes in Auditory Category Learning

Casey L. Roark[1,2] **(croark@pitt.edu)**

Bharath Chandrasekaran[1,2] **(b.chandra@pitt.edu)**

[1]Communication Science and Disorders, University of Pittsburgh, Pittsburgh, PA 15213 USA
[2]Center for the Neural Basis of Cognition, Pittsburgh, PA, 15213 USA

Abstract

Learning the sounds of a new language depends on the ability to learn novel auditory categories. Multidimensional categories, whether speech or nonspeech, can be learned through feedback and different category structures are proposed to recruit separate cognitive and neural mechanisms. There is substantial individual variability in learning; however, it is rare to compare learning of different categories in the same individuals. Understanding the sources of variability has theoretical implications for category learning. In this study, we trained the same participants on three types of multidimensional auditory categories. Participants learned nonspeech rule-based, nonspeech information-integration, and Mandarin speech categories. Learning of all category types was related across individuals and differences in working memory similarly supported learning across tasks. There was substantial variability in learning outcomes and strategies used to learn the categories. There are multiple paths to successful learning and appreciation of individual variability is essential to understanding the mechanisms of learning.

Keywords: category learning; audition; working memory; individual differences

Introduction

Complex natural categories, such as speech sound categories, are often multidimensional (Hillenbrand et al., 1995). Learning the sounds of a new language requires mapping these multidimensional inputs into novel categories (Holt & Lotto, 2010). The structure of categories in multidimensional space has been proposed to be a key factor driving the cognitive and neural mechanisms that support successful learning (Ashby et al., 1998; Ashby & Maddox, 2011).

Specifically, according to a dual systems framework initially proposed by Ashby et al. (1998), at least two systems support category learning. Originally developed in the visual modality, this theory has recently been expanded to the auditory modality (Chandrasekaran, Yi & Maddox, 2014). The dual systems framework proposes that an *explicit* system supports learning of categories that can be described as *rule-based* (RB) categories. RB category learning requires selective attention to individual stimulus dimensions and relies on hypothesis generation and testing (Ashby et al., 1998). In contrast, the *implicit* system supports optimal learning of *information-integration* (II) categories. II category learning requires pre-decisional integration along multiple stimulus dimensions and relies on procedural learning processes (Ashby et al., 1998). RB and II categories can both be multidimensional. A key feature that is thought to differentiate the two types of categories is that RB categories can be distinguished by rules that are easy to describe verbally and II categories are difficult to describe.

Typically, RB and II categories studied in the lab are artificial categories, generated by experimenters to examine novel learning. However, it has also been proposed that natural categories like speech categories can be thought of as II categories (Yi et al., 2016). Many of the world's languages are *tonal* languages, meaning that pitch-related information of a syllable or word is lexically distinctive. For example, in Mandarin, the same syllable (e.g., /ma/) produced with four different tones has four unique meanings. Mandarin tone categories vary across multiple pitch-related dimensions, cannot easily be described by speakers of non-tonal languages and, thus, can be difficult to learn. Research has demonstrated that Mandarin tone learning may rely on similar neural mechanisms as II learning (Yi et al., 2016) and is affected by the same manipulations that affect II learning (Chandrasekaran et al., 2014; Maddox et al., 2014; but see Maddox et al., 2013). However, learning of Mandarin and artificial II or RB categories has never been compared in the same individuals.

Even in the same task, there is often a wide variability in performance (e.g., Llanos et al., 2020; Shamloo & Hélie, 2020) – while some individuals learn the categories quickly and quite well, others struggle to learn. Studies of RB and II learning tend to focus on the differences between the tasks, rather than comparing how the same individuals approach the tasks. To better understand the sources of individual variability during learning, we compare final block accuracies of the same individuals across tasks as measures of learning outcomes. One possibility is that individuals who succeed in one task may also succeed in the other tasks. However, if Mandarin is more similar to an II problem, II and Mandarin learning may be more strongly related to one another than to RB learning.

Individual differences in cognitive abilities have also been proposed to differentiate RB and II learning. Specifically, the role of working memory (WM) has been proposed to be vitally important for RB learning and either minimally or not involved in II learning (Ashby et al., 1998). The evidence for a differential role of WM in II and RB learning is quite mixed (DeCaro et al., 2008; Lewandowsky et al., 2012; Miles & Minda, 2011; Miles, Matsuki, & Minda, 2014; Newell et al., 2013; Newell, Dunn, & Kalish, 2011; Roark & Holt, 2019; Zeithamova & Maddox, 2007). A competing perspective argues that WM supports successful learning, regardless of the underlying category structure (Kalish, Newell, & Dunn, 2017). In this study, we examine how individual differences in WM are related to learning outcomes. If, as the dual systems perspective predicts, WM is more relevant for successful RB learning compared to II or Mandarin tone learning, we should see a stronger relation between WM and RB learning outcomes than II or Mandarin learning outcomes. If instead, WM supports successful learning regardless of category structure, we should see a similar relation between WM and learning outcomes of all tasks. Assessing these alternative hypotheses in the same individuals will provide a better understanding of the role of WM in learning categories with different structures.

Finally, we compare the types of strategies individuals use to learn the categories. During learning, individuals can use different strategies to separate stimuli into categories. They can separate the categories optimally based on the component dimensions and attending only to relevant information. Alternatively, they may use suboptimal strategies, attending to the underlying dimensions in a way that is not aligned with the task. Individual differences in strategies determine what information is encoded and remembered and is related to how well an individual learns. The strategies that individuals use across different tasks has not been explored. It is possible that successful learners are successful because they can flexibly shift their strategies based on what is optimal for the task. However, it is also possible that individuals may be successful despite using suboptimal strategies.

Methods

Across two sessions separated by at least one week, participants learned nonspeech RB, nonspeech II, and Mandarin speech categories. The order of the two nonspeech tasks was counterbalanced across sessions. After the first task, participants completed an automated version of the operation span task as a measure of WM capacity.

Participants

Participants were recruited through Prolific (www.prolific.sc) and the experiment was administered using the online Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2019). One hundred participants ages 18-35 (45 F, *M* 25.3 years, *sd* 5.05 years) completed one session and 90 returned for a second session. Of these 90 participants, 86 (36 F, *M* 25.4 years, *sd* 5.04 years) completed all three tasks and four did not complete the Mandarin task. The analyses included participants who completed all three tasks. All participants were native speakers of non-tonal languages and reported no prior experience with any tonal languages including Mandarin. Participants received \$10/hour for their participation for a total of \$20 across two sessions.

Stimuli

All three category types required participants to use multidimensional information to distinguish four possible sound categories. Stimuli for the nonspeech tasks were nonspeech ripples varying in temporal modulation and spectral modulation (Fig. 1A-B). These dimensions are thought to be fundamental properties of complex sounds, including speech (Woolley et al., 2005). The nonspeech categories were created by sampling from a bivariate normal distribution and had 300 total stimuli (75 stimuli/category).

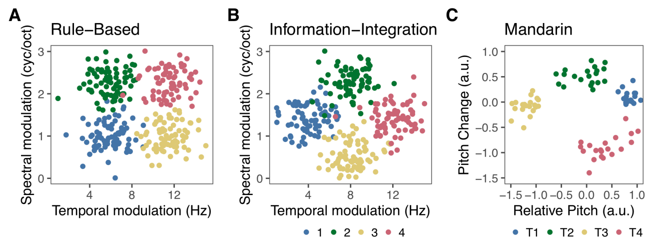

Figure 1: **A.** Nonspeech Rule-Based and **B.** Information-Integration category structures varying in temporal modulation (Hz) and spectral modulation (cyc/oct). **C.** Mandarin tone (T) speech categories varying in relative pitch (arbitrary units) and pitch change (arbitrary units).

Stimuli for the Mandarin category learning task were natural speech productions recorded from native speakers of Mandarin Chinese and varying along relative pitch and pitch change dimensions (Fig. 1C). The lexical tone stimuli were produced in five syllable contexts (/bu/, /di/, /lu/, /ma/, and /mi/) by four speakers (2F, 2M) for a total of 80 stimuli (20 stimuli/category). The stimuli from two speakers (1F, 1M) were used during training and the other two speakers (1F, 1M) were used in the generalization test. The stimuli are clearly perceived as speech and were duration-normalized to 440 ms and RMS amplitude matched to 70 dB.

Procedure

Participants completed two sessions, separated by at least one week. In session 1, participants completed one of the nonspeech learning tasks (RB or II) followed by an automated version of the operation span task (OSPAN, Unsworth et al., 2005) as a measure of WM capacity. In session 2, participants completed the other nonspeech learning task followed by the Mandarin tone learning task. Before the first task in each session, participants completed a sound check to ensure that they could hear the sounds and that they were wearing headphones. Specifically, three 1000 Hz sounds played with an ISI of 100 ms in one ear at a time. Participants reported in which ears the sounds played.

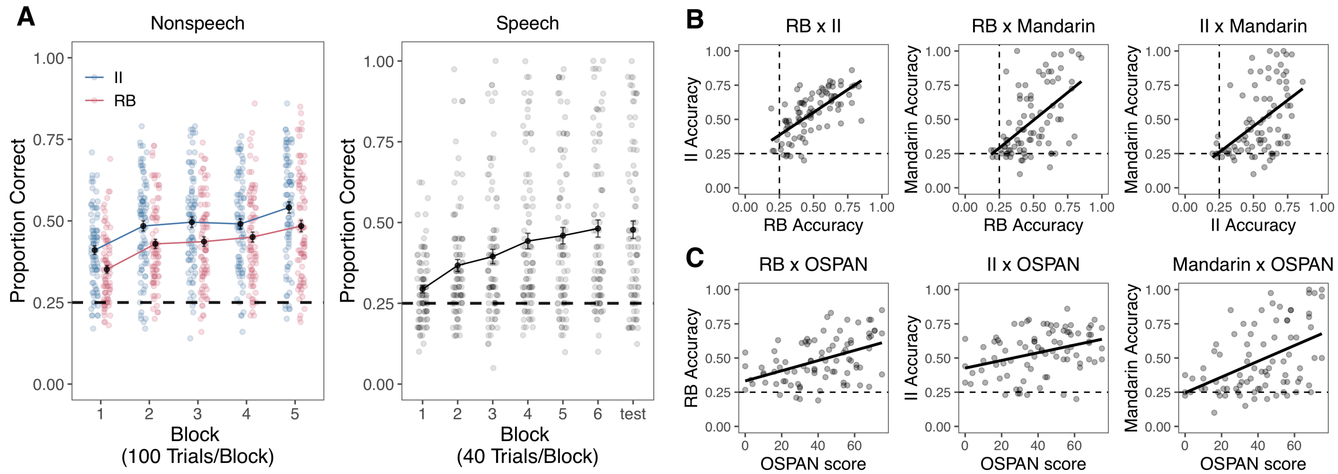

Figure 2: **A.** Block-by-block accuracy for the nonspeech Rule-Based (RB) and Information-Integration (II) tasks and Mandarin speech learning task relative to chance performance (25%) shown as a dashed-line. Lines reflect the mean and error bars reflect *SEM.* **B.** Correlations between final block accuracy in all three tasks individually. **C.** Correlations between OSPAN working memory score and final block accuracy in all three tasks individually.

Category Learning Participants were instructed to learn the four equally likely categories by listening to the sounds and using feedback to become more accurate. Participants learned the RB and II categories across five 100-trial blocks and the Mandarin categories across six 40-trial blocks. There were fewer trials in the Mandarin task because there are fewer category exemplars. Stimuli were presented dichotically for a duration of 1 sec (II/RB) or 440 ms (Mandarin). Participants pressed one of four buttons (1, 2, 3, or 4) and received feedback ("Correct"/"Incorrect") for 1 sec followed by a 1 sec ITI. After Mandarin training, participants categorized 40 speech tokens from two novel speakers (1M, 1F) and no longer received feedback.

Working Memory To assess WM capacity, participants completed an automated version of the operation span task (OSPAN, Unsworth et al., 2005). Participants were shown simple arithmetic problems (e.g., (9 + 5) x 1 = 15), reported whether the presented solutions were correct or incorrect, and were then shown a letter on the screen (e.g., W). After a full sequence was presented, participants recalled the letters presented in order. Participants saw 15 letter sequences that spanned from three to seven letters. The OSPAN score was calculated as the sum of the length of all of the correctly recalled spans. For instance, if a participant correctly recalled the sequence of four letters, four points were added to their score. We did not filter the scores on the basis of accuracy (Đokić, Koso-Drljević, & Đapo, 2018).

Results

We examined how learning outcomes were related across tasks, how WM capacity is related to learning outcomes, and the strategies participants used within and across tasks.

Learning within and across tasks

Participants were successful at both nonspeech and speech learning, with substantial variability across individuals (Fig. 2A). We ran a 2 (II, RB) x 5 (block) repeated measures ANOVA to assess differences in II and RB learning. Accuracy generally improved across blocks ($F(1, 340) = 67.7$, $p < .001$, $\eta_p^2 = .44$) and was higher for the II task ($M = 49\%$) than the RB task ($M = 43\%$; $F(1, 85) = 27.8$, $p < .001$, $\eta_p^2 = .25$). There was no interaction between block and category ($F(4, 340) = 0.54$, $p = .71$, $\eta_p^2 = 0.006$).

To assess learning in the speech task, we ran a 6-level repeated measures ANOVA to assess learning across blocks. Performance generally improved across blocks ($F(5, 425) = 35.5$, $p < .001$, $\eta_p^2 = .30$). Participants were also able to generalize their category knowledge to two novel speakers (*M:* 48%, one-sample $t(85) = 8.45$, $p < .001$, $d = 0.91$).

After confirming that participants learned the categories and that there was substantial individual variability in learning, we compared how learning outcomes were related across tasks by examining the correlations between final-block accuracies across tasks (Fig. 2B).

Final block accuracy was significantly positively correlated across all tasks (RB-II: $r(84) = .67$, $p < .001$; RB-Mandarin: $r(84) = .54$, $p < .001$; II-Mandarin: $r(84) = .50$, $p < .001$). We compared the strengths of these correlations using the *cocor* package in R (Diedenhofen & Musch, 2015). There were no significant differences in the correlations between RB-Mandarin and II-Mandarin ($z = 0.47$, $p = .64$) or RB-II and RB-Mandarin ($z = 1.63$, $p = .10$). However, the correlation between II-RB was significantly stronger than the correlation between II-Mandarin ($z = 2.08$, $p = .037$).

Working memory and learning outcomes

A separate question regards the association between WM capacity (indexed by the OSPAN score) and learning outcomes in each of the three tasks (Fig. 2C).

WM capacity was significantly positively correlated with accuracy in all three tasks (RB: $r(84) = .46$, $p < .001$, II: $r(84) = .35$, $p < .001$, Mandarin: $r(84) = .47$, $p < .001$). There were no significant differences in the correlations of any task's

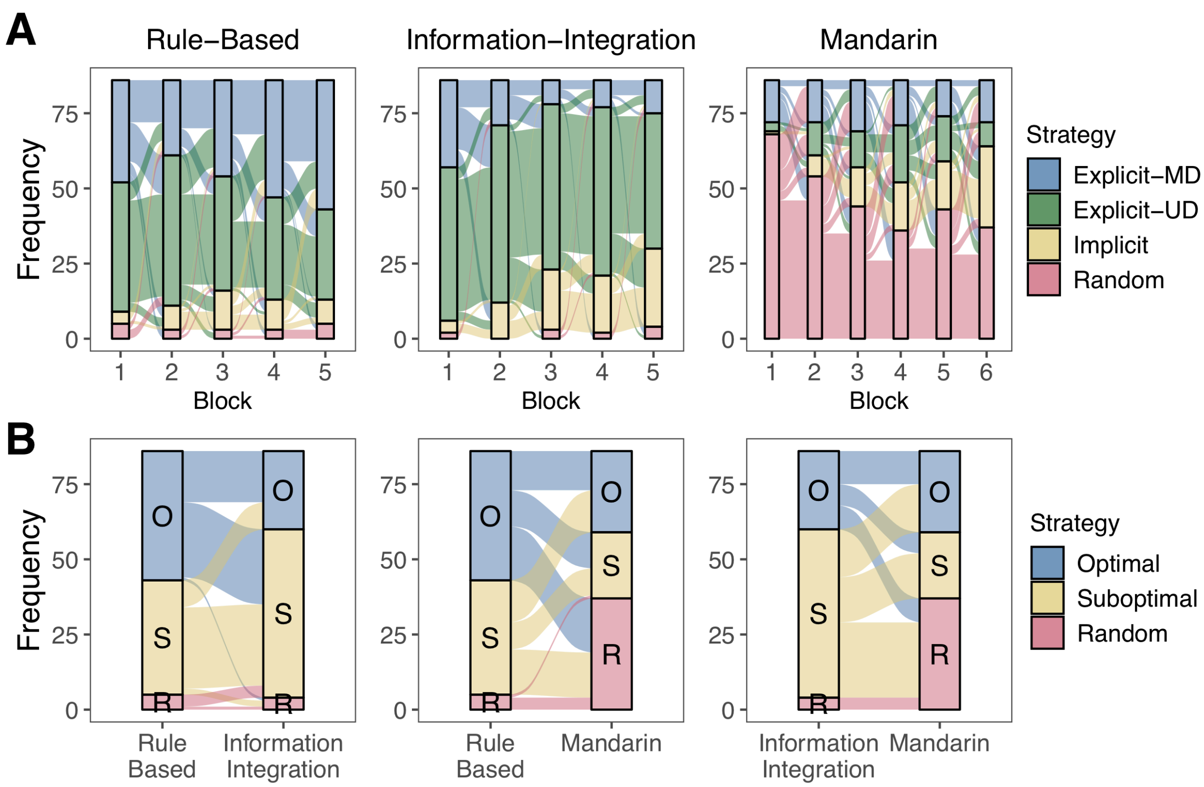

Figure 3: **A.** Strategies across all blocks with the movement of participants between strategies shown in shaded regions. **B.** Strategy type (Optimal/Suboptimal/Random) in the final block of each of the three tasks. The Optimal strategy was Explicit-MD for the Rule-Based task and Implicit for the Information-Integration and Mandarin tasks.

final block accuracy and WM capacity (WM-II vs. WM-RB: $z = 1.37$, $p = .17$; WM-II vs. WM-Mandarin: $z = 1.24$, $p = .21$; WM-RB vs. WM-Mandarin: $z = 0.12$, $p = .91$). WM was not differently associated with RB, II, or speech learning.

Strategies during learning

To understand the strategies individuals use when learning different kinds of categories, we applied several classes of decision-bound computational models (Ashby, 1992; Maddox & Ashby, 1993). These models assume that participants divide the categories in the two-dimensional space with decision boundaries that can rely on implicit or explicit learning processes. We fit classes of models that make different assumptions: implicit model, explicit multidimensional and unidimensional models, and a random responder model. We chose these models because they allow for assessment of different dimension-based strategies during RB and II learning (Ashby & Maddox, 2011).

The *implicit* Striatal Pattern Classifier reflects procedural learning mechanisms and assumes that participants use feedback to learn stimulus-response associations (Ashby et al., 1998; Ashby & Waldron, 1999) and can be thought of as a complex implementation of an exemplar model (Ashby & Rosedahl, 2017). This model has nine free parameters: eight that determine the location of hypothetical 'striatal' units in perceptual space and one that represents the noise associated with the placement of the units.

The second class of models represents *explicit*, hypothesis-testing mechanisms and includes multidimensional (MD) and unidimensional (UD) models. Explicit-MD models assume that the participant places two decision boundaries (one along each dimension) that are combined to determine category membership and have three free parameters: two for the boundaries along the x and y-dimensions and one noise parameter. Explicit-UD models assume that the participant sets three decision boundaries along only one of the dimensions (x-dimension or y-dimension) and have four free parameters: three for boundaries along the relevant dimension and one noise parameter.

The random responder model that assumes that the participant guesses on each trial.

Model parameters were estimated using maximum likelihood procedures (Wickens, 1982) and model selection used the Bayesian Information Criterion (BIC), which penalizes models with more free parameters: BIC = $r{*}\ln N$ - $2\ln L$, where r is the number of free parameters, N is the number of trials in a given block for a given participant, and L is the likelihood of the model given the data (Schwarz, 1978). For each block for each participant, the model with the lowest BIC value was selected as the best-fitting model.

Participants can use different kinds of strategies and be successful in these tasks. However, each category type has an Optimal strategy that is particularly suited for learning and allows participants to achieve the highest possible accuracies.

All other kinds of strategies are Suboptimal. Suboptimal strategies (except for Random strategies) may still lead to above-chance learning but are not associated with the highest accuracy. The Implicit strategy is Optimal for the II and Mandarin tone categories and the Explicit-MD strategy is Optimal for the RB categories.

We first examined the strategies participants used *within* the three tasks (Fig. 3A). Within each task, participants used a variety of strategies. The most common strategies were the Optimal, Explicit-MD strategy for the RB task, the Suboptimal, Explicit-UD strategy for the II task, and a Random strategy for the Mandarin task. Within each task, participants commonly switched between different strategies across blocks (RB: $M = 1.86$, $Mdn = 2$ switches; II: $M = 1.35$, $Mdn = 1$; Mandarin $M = 1.86$, $Mdn = 2$ across the first five blocks). The number of switches was statistically different across the tasks (non-parametric Friedman test, $\chi^2(2) = 8.60$, $p = .014$, $W = 0.050$). Bonferroni-corrected pairwise Wilcoxon signed rank tests indicated that there were significantly more strategy switches in the Mandarin task than II task ($p = .032$), but no differences between the II and RB tasks ($p = .11$) or RB and Mandarin tasks ($p = 1.0$).

We next examined the strategies in the final block *across* the three tasks (Fig. 3B) and grouped strategies by whether they were Optimal, Suboptimal, or Random. As expected, participants had higher accuracies when they used Optimal compared to Suboptimal strategies across all three tasks (RB: Optimal 53%, Suboptimal 46%, $t(70.5) = 2.35$, $p = .021$, $d = 0.51$; II: Optimal 68%, Suboptimal 50%, $t(66.8) = 7.00$, $p < .001$, $d = 1.47$; Mandarin: Optimal 77%, Suboptimal 45%, $t(43.8) = 6.68$, $p < .001$, $d = 1.93$). Assessing strategies across tasks gives a better understanding of how participants are able to *flexibly* shift their strategies based on the demands of the tasks. If participants are able to flexibly shift their attention, they will use an Optimal strategy across tasks. If instead, they have difficulty shifting their attention, they may use Suboptimal strategies or a mix of Suboptimal and Optimal strategies across tasks.

For the RB and II tasks, 20% of participants used an Optimal strategy in both nonspeech tasks, 32% used a Suboptimal strategy in both tasks, and 40% used a Suboptimal strategy in one task and an Optimal strategy in the other task. The remaining 9% used a Random strategy in at least one task. This does not sum to 100% because of rounding of percentages.

For the RB and Mandarin tasks, 15% used an Optimal strategy in both tasks, 10% used a Suboptimal strategy in both tasks, and 30% used a Suboptimal strategy in one task but an Optimal strategy in the other task. The final 44% used a Random strategy in at least one task. This does not sum to 100% because of rounding of percentages.

Unlike the other comparisons, the Optimal strategy for II and Mandarin tasks is the same Implicit strategy. Among all participants, 13% used an Optimal strategy in both II and Mandarin tasks, 17% used a Suboptimal strategy in both tasks, and 27% used a Suboptimal strategy in one task but an Optimal strategy in the other task. The other 43% participants used a random strategy in at least one of the tasks.

The large number of participants using a Random strategy in the Mandarin task may be related to the poorer performance in this task. Participants using a Random strategy had an average accuracy of 29%, which is near chance (25%). In contrast, participants using an Optimal strategy had an average accuracy of 77%. It is not the case that individuals cannot learn these categories or that there is an inherent difficulty in the ability of the models to account for the data. Rather, many participants struggled to learn.

These results demonstrate that participants use a variety of strategies to learn auditory categories. While accuracy in one task is correlated with accuracy in the other tasks, participants do not necessarily use Optimal strategies across tasks to achieve that performance.

Discussion

We compared learning of three types of multidimensional auditory categories in the same individuals. Performance was related across tasks, but there was substantial variability in learning outcomes and strategies. Individual differences in WM were related to individual differences in learning outcomes to similar extents in all three tasks. These results highlight the importance of consideration of individual differences in learning, both within and between tasks.

Generally, if a participant succeeded in one task, they succeeded in the other tasks. However, we found that performance in the two nonspeech tasks was more strongly correlated than performance in the II and Mandarin tasks. Mandarin tone categories have been proposed to be a type of II task based on similarities in their structure (Yi et al., 2016; Fig. 1). However, our results indicate that other similarities between tasks are also important for understanding how an individual will learn. The dimensions or nonspeech nature of the RB and II tasks may have led them to be more similar to one another than the II and Mandarin tasks. Further, many participants struggled to learn the Mandarin categories. This indicates that components other than the category structure affect learning outcomes. For these speech stimuli, information about talker or syllable may have factored into participants' categorization decisions. In all, elements that support similarities in learning outcomes across tasks must be considered.

Understanding the role of individual differences in cognitive abilities, like WM, is important for understanding the mechanisms that support learning. We found that WM was similarly positively correlated with learning outcomes in all three tasks. This finding is counter to the proposal that WM supports successful learning of RB, but not II, categories (Ashby et al., 1998) and instead aligns with the perspective that WM supports successful learning, regardless of category structure (Kalish et al., 2017; Lewandowsky et al., 2012). While our results do not differentiate between single system or dual systems accounts, they suggest a role for individual differences in WM contributing to individual differences in learning outcomes, regardless of category structure.

Finally, examining the strategies participants use during category learning provides a novel perspective on individual variability during learning. While some individuals flexibly shifted towards Optimal strategies across different tasks, many individuals used an Optimal strategy in one task and a Suboptimal strategy in another. There are multiple paths to success during learning and consideration of individual differences in strategy use is critical for understanding individual differences in category learning. Theories of category learning should focus on explaining the factors that determine whether individual will be successful in category learning, demonstrate selective advantages, or be unsuccessful in learning regardless of category structure.

In sum, there is substantial variability in learning outcomes during auditory category learning. To understand what drives the individual variability in learning, we examined learning of three types of auditory categories in the same individuals. Learning outcomes for nonspeech RB, nonspeech II, and Mandarin speech categories were all positively correlated. WM capacity was positively correlated with learning outcomes across all three tasks, with no differences among tasks. Finally, individuals demonstrated substantial variability in the strategies they used to learn the categories. These results highlight the importance of consideration of the sources of individual variability in category learning.

Acknowledgments

This research was supported by the National Institute on Deafness and Other Communication Disorders (R01DC013315A1 to B.C. and F32DC018979 to C.L.R.). This work used the Extreme Science and Engineering Discovery Environment (XSEDE, Towns et al., 2014), which is supported by NSF (ACI-1548562). Specifically, it used the Bridges system (Nystrom et al., 2015) which is supported by NSF (ACI-1445606) at the Pittsburgh Supercomputing Center.

References

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. (2019). Gorilla in our Midst: An online behavioral experiment builder. *Behavior Research Methods*, 438242.

Ashby, F. G. (1992). *Multidimensional models of categorization* (F. G. Ashby, Ed.; pp. 449–483). Lawrence Erlbaum.

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*(3), 442–481.

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147–161.

Ashby, F. G., & Rosedahl, L. (2017). A Neural Interpretation of Exemplar Theory. *Psychological Review*, *124*(4), 472–482.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review*, *6*(3), 363–378.

Chandrasekaran, B., Yi, H.-G., & Maddox, W. T. (2014). Dual-learning systems during speech category learning. *Psychonomic Bulletin & Review*, *21*, 488–495.

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.

Diedenhofen, B. & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLOS ONE*, *10*(4), e0121945.

Đokić, R., Koso-Drljević, M., & Đapo, N. (2018). Working memory span tasks: Group administration and omitting accuracy criterion do not change metric characteristics. *PLOS ONE*, *13*(10), e0205169.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97*(5 Pt 1), 3099–3111.

Holt, L. L. & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception, & Psychophysics*, *72*(5), 1218–1227.

Kalish, M. L., Newell, B. R., & Dunn, J. C. (2017). More Is Generally Better: Higher Working Memory Capacity Does Not Impair Perceptual Category Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 503–514.

Lewandowsky, S., Yang, L.-X., Newell, B. R., & Kalish, M. L. (2012). Working memory does not dissociate between different perceptual categorization tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 881–904.

Llanos, F., McHaney, J. R., Schuerman, W. L., Yi, H. G., Leonard, M. K., & Chandrasekaran, B. (2020). Non-invasive peripheral nerve stimulation selectively enhances speech category learning in adults. *Npj Science of Learning*, *5*(1), 12.

Maddox, W. T. & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*(1), 49–70.

Maddox, W. T., Chandrasekaran, B., Smayda, K., & Yi, H.-G. (2013). Dual systems of speech category learning across the lifespan. *Psychology and Aging*, *28*(4), 1042–1056.

Maddox, W. T., Chandrasekaran, B., Smayda, K., Yi, H.-G., Koslov, S., & Beevers, C. G. (2014). Elevated depressive symptoms enhance reflexive but not reflective auditory category learning. *Cortex*, *58*, 186–198.

Miles, S. J., Matsuki, K., & Minda, J. P. (2014). Continuous executive function disruption interferes with application of an information integration categorization strategy. *Attention, Perception, & Psychophysics*, *76*(5), 1318–1334.

Miles, S. J. & Minda, J. P. (2011). The effects of concurrent verbal and visual tasks on category learning. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *37*(3), 588–607.

Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of Category Learning Fact or Fantasy? *Psychology of Learning and Motivation*, *54*, 167–215.

Newell, B. R., Moore, C. P., Wills, A. J., & Milton, F. (2013). Reinstating the Frontal Lobes? Having More Time to Think Improves Implicit Perceptual Categorization: A Comment on Filoteo, Lauritzen, and Maddox (2010). *Psychological Science*, *24*(3), 386–389.

Nystrom, N. A., Levine, M. J., Roskies, R. Z., and Scott, J. R. (2015). Bridges: A Uniquely Flexible HPC Resource for New Communities and Data Analytics. In *Proceedings of the 2015 Annual Conference on Extreme Science and Engineering Discovery Environment* (St. Louis, MO, July 26-30, 2015). XSEDE15. ACM, New York, NY, USA.

Roark, C. L. & Holt, L. L. (2019). Auditory information-integration category learning in young children and adults. *Journal of Experimental Child Psychology*, *188*, 104673.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

Shamloo, F. & Hélie, S. (2020). A study of individual differences in categorization with redundancy. *Journal of Mathematical Psychology*, *99*, 102467.

Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G.D., Roskies, R., Scott, J.R. and Wilkens-Diehr, N. (2014). XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering*, 16(5):62-74.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505.

Wickens, T. D. (1982). *Models for Behavior: Stochastic Processes in Psychology*. W. H. Freeman.

Woolley, S. M. N., Fremouw, T. E., Hsu, A., & Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature Neuroscience*, *8*(10), 1371–1379.

Yi, H.-G., Maddox, W. T., Mumford, J. A., & Chandrasekaran, B. (2016). The Role of Corticostriatal Systems in Speech Category Learning. *Cerebral Cortex*, *26*(4), 1409–1420.

Zeithamova, D. & Maddox, W. T. (2007). The role of visuospatial and verbal working memory in perceptual category learning. *Memory & Cognition*, *35*(6), 1380–1398.