Published in final edited form as:

J Exp Child Psychol. 2016 February ; 142: 48–65. doi:10.1016/j.jecp.2015.09.018.

The Role of Age and Executive Function in Auditory Category Learning

Rachel Reetzke¹, W. Todd Maddox^{2,3,4,5}, and Bharath Chandrasekaran^{1,2,3,4}

¹Department of Communication Sciences and Disorders, Moody College of Communication, The University of Texas at Austin

²Department of Psychology, College of Liberal Arts, The University of Texas at Austin

³Institute of Mental Health Research, College of Liberal Arts, The University of Texas at Austin

⁴The Institute for Neuroscience, College of Liberal Arts, The University of Texas at Austin

⁵Center for Perceptual Systems, College of Liberal Arts, The University of Texas at Austin

Abstract

Auditory categorization is a natural and adaptive process that allows for the organization of high-dimensional, continuous acoustic information into discrete representations. Studies in the visual domain have identified a rule-based learning system that learns and reasons via a hypothesis-testing process that requires working memory and executive attention. The rule-based learning system in vision shows a protracted development, reflecting the influence of maturing prefrontal function on visual categorization. The aim of the current study is two-fold: (a) to examine the developmental trajectory of rule-based auditory category learning from childhood through adolescence, into early adulthood; and (b) to examine the extent to which individual differences in rule-based category learning relate to individual differences in executive function. Sixty participants with normal hearing, 20 children (age range, 7–12), 21 adolescents (age range, 13–19), and 19 young adults (age range, 20–23), learned to categorize novel dynamic ripple sounds using trial-by-trial feedback. The spectrotemporally modulated ripple sounds are considered the auditory equivalent of the well-studied Gabor patches in the visual domain. Results revealed that auditory categorization accuracy improved with age, with young adults outperforming children and adolescents. Computational modeling analyses indicated that the use of the task-optimal strategy (i.e. a conjunctive rule-based learning strategy) improved with age. Notably, individual differences in executive flexibility significantly predicted auditory category learning success. The current findings demonstrate a protracted development of rule-based auditory categorization. The results further suggest that executive flexibility coupled with perceptual processes play important roles in successful rule-based auditory category learning.

Correspondence: Bharath Chandrasekaran, The University of Texas at Austin, Department of Communication Sciences and Disorders, 2504A Whitis Ave., Austin, TX 78712, Phone: 512-471-4119, Fax: 512-471-2957, bchandra@utexas.edu.

Conflicts of Interest: None

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

development; category learning; rule-based learning; executive function; perceptual processes

Introduction

Auditory category learning is a fundamental cognitive process that allows for the organization of high-dimensional, continuous acoustic information into discrete representations (Gifford, Cohen, & Stocker, 2014). Whether it is speech, musical scales, auscultation, or sound patterns (e.g. Morse code), learning to categorize new auditory information progresses throughout the lifespan. The ability to learn natural speech and musical categories is well studied in infancy and adulthood (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Chandrasekaran, Yi, & Maddox, 2014; Kuhl, 2004; Trainor & Trehub, 1992; Yi, Maddox, Mumford, & Chandrasekaran, 2014). To our knowledge, there has been little attempt to *compare* the extent to which children and adults learn novel auditory categories. The current study focuses on tracking the development of rule-based (RB) auditory category learning using artificial, experimenter-constrained category structures.

Previous work has shown that speech and musical categories are easily acquired in early infancy (Kuhl, 2004; Trainor & Trehub, 1992). The consensus is that speech and musical categories are optimally acquired via implicit learning processes (Kuhl, 2004) that are dependent on early maturing neural circuitry (Chandrasekaran, Yi, Blanco, McGeary, & Maddox, 2015; Yi et al., 2014). Learning speech and musical categories in adulthood however, is known to be a more laborious process, requiring extensive auditory training (Lively, Logan, & Pisoni, 1993; Wong, Morgan-Short, Ettlinger, & Zheng, 2012). Several examples come from the speech domain: infants can rapidly acquire fine-grained speech categories within the first year of life (Kuhl, 2004). For example, infants exposed to English are able to distinguish between the phonemes /l/ and /r/ within the first year of life, both of which are phonemic in English (Kuhl, 2004). In contrast, Japanese-speaking adults struggle to learn the /l/ and /r/ category distinction, which are not phonemic in Japanese, even after extensive auditory training (Lively et al., 1993). These studies suggest fundamental differences in learning capacities as a function of age, with a potential disadvantage for acquiring speech categories later in life (Wong et al., 2012). Whether the general finding that poorer auditory category learning in adults relative to children generalizes to rule-based auditory categories has not been examined.

In the visual domain, extensive neuropsychological, neuroimaging, behavioral and computational modeling studies have identified a fast-mapping and slow maturing rule-based, hypothesis-testing learning system, *in addition* to a slow-mapping and earlier maturing implicit, procedural-based learning system (Ashby, Paul, & Maddox, 2011; Ashby & Maddox, 2011). Dual-learning systems theoretical framework, like the COVIS (Competition between Verbal and Implicit Systems), posit that the two learning systems are at least partially dissociable, with an early dominance for the rule-based, hypothesis testing system in humans (Ashby et al., 2011; Ashby & Maddox, 2011). The rule-based system is

dependent on working memory and executive attention to generate and maintain rules, based on the visual dimensions underlying the stimuli. In contrast, the implicit, procedural-based learning system is not under conscious control (Ashby & Maddox, 2011). In this form of learning, striatal neurons learn to associate a group of sensory units with a motor response when learning is rewarded (Ashby & Maddox, 2011). The systems are neurally dissociable. The implicit learning system is critically-dependent on the striatum, and is likely to mature early in infancy (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Nomura et al., 2007; Seger, 2008). In contrast, the rule-based, hypothesis testing system is predominantly driven by the prefrontal cortex, which shows a protracted maturational profile (Ashby et al., 1998; Gabrieli, Brewer, Desmond, & Glover, 1997; Schacter & Wagner, 1999). The PFC is not completely mature until early adulthood. Indeed, in studies examining rule-based learning in the visual domain, a consistent finding is that adults outperform children, especially when categorization involves complex rules. This performance difference has been attributed to protracted development of executive function in children (Casey, Tottenham, Liston, & Durston, 2005; Crone, Donohue, Honomichl, Wendelken, & Bunge, 2006; Huang-Pollock, Maddox, & Karalunas, 2011; Minda, Desroches, & Church, 2008; Rabi, Miles, & Minda, 2015; Rabi & Minda, 2014). Rabi and Minda (2014) examined the learning of a rule-based task in children (4–11) and adults (18–27). The stimuli in their experiment were Gabor patches that varied in spatial frequency and orientation. For this two-category learning task, the computational strategy that maximized accuracy was to place low spatial frequency items into category A and high spatial frequency items into category B, while ignoring the irrelevant variation in spatial orientation. The results revealed that: a) rule-based learning shows a protracted development, as indexed by learning accuracies and the use of the optimal strategy, b) performance across groups was equivalent when comparing adults and children who used the optimal strategy, and c) within the children, working memory capacity, was related to improvements in rule-based category learning as a function of age.

The extent to which results from studying rule-based learning in the visual domain can be generalized to the auditory domain is an open question. There is clearly a human capacity to learn auditory categories via rules that is sustained even in adulthood (Bradlow et al., 1997; Chandrasekaran, Koslov, & Maddox, 2014; Chandrasekaran, Yi, et al., 2014). Medical students and emergency medical technicians are taught (in adulthood) to interpret auscultation on the basis of rules. For example, a high-pitched, discontinuous crackling lung sound during inspiration is indicative of abnormality (therefore critical in normal versus abnormal lung auscultation categorization). These auditory rules can be complex and multidimensional—when the ‘crackle’ appears early in the inspiratory phase, it can be indicative of bronchitis; late in the inspiratory phase it can indicate pneumonia. Rules can be constructed using simple auditory dimensions as well (Lichtenstein et al., 2004). The International Morse Code, for example, uses rules based on simple standardized short and long auditory signals called “dots” (a 50 ms tone) and “dashes” (made up of three ‘dots’) to communicate complex information. Speech and musical categories, the most ubiquitous auditory signals, are optimally learned implicitly, but can also be acquired (likely suboptimally) via explicit, instructional methods (Maddox & Chandrasekaran, 2014). For example, in Mandarin, a tone language, pitch trajectories within a syllable can indicate different words, and hence are a categorical distinction in that language. Second language

learners are explicitly taught these category distinctions using rules: tone 1 has a high-level pitch trajectory, tone 2 has a low-rising pitch trajectory; tone 3 has a low-dipping pitch trajectory; tone 4 has a high-falling pitch trajectory.

The concept of multiple learning systems mediating categorization has been recently applied in the auditory domain, in studying vowel processing (Maddox, Molis, & Diehl, 2002), as well as learning tonal categories in adulthood (Chandrasekaran, Koslov, et al., 2014; Chandrasekaran, Yi, et al., 2014). These studies have shown that successful learning of speech categories activates the neural circuitry involved in implicit, procedural-learning (Yi et al., 2014). However, the circuitry underlying the rule-based system, including the PFC is also activated during learning, but is not associated with successful learning (Yi et al., 2014). Computational modeling results indicate that successful speech learning requires switching control from the rule-based system to the procedural-based system (Chandrasekaran et al., 2015; Chandrasekaran, Yi, et al., 2014; Maddox, Chandrasekaran, Smayda, & Yi, 2013; Maddox et al., 2014; Yi et al., 2014). However, disambiguating the relative contribution of the two learning systems is challenging with natural, unconstrained speech categories. For example, similar learning accuracies can be achieved by using multidimensional, conjunctive strategies, as well as multidimensional implicit, procedural-based learning strategies (Maddox & Chandrasekaran, 2014). There is therefore a need to study the learning of novel auditory category structures that are constrained by the experimenter, so that the optimal strategy can be predetermined (for example, to be rule-based).

Current study

In the current study, we examine the developmental trajectory of auditory rule-based category learning by studying children, adolescents, and adults using behavioral and computational modeling methods. In addition, we examine the extent to which conjunctive, rule-based strategies are associated with individual differences in executive function, independent of age. This study addresses a number of unanswered questions and will ultimately shed light on the underlying perceptual and cognitive factors associated with optimal auditory rule-based category learning. In the next few sections, we expand on our experimental design by focusing on the motivation for our choice of category structure and auditory dimensions. We will then discuss the computational modeling methods and conclude with specific hypotheses.

Category structure—We use a two-dimensional, four-category rule-based structure modeled after prior work in vision. In the visual domain, studies have shown that this type of multidimensional category structure (shown in Figure 1) is impacted by training manipulations that target the rule-based learning system, but not those that target the procedural-based learning system (Maddox & David, 2005; Maddox, Filoteo, & Lauritzen, 2007; Maddox, Love, Glass, & Filoteo, 2008). The advantage of using multidimensional, conjunctive rule-based categories is that they are more likely to yield a broader age effect (because they are more challenging to learn). In addition, it is likely that the use of multidimensional (i.e. conjunctive rules) relative to unidimensional rules will increase with

age, and we will be able to test this rigorously for the first time through computational modeling.

Auditory dimensions—As shown in Figure 1, the two dimensions used are spectral and temporal modulation. There are several motivations for the choice of dimensions. 1) spectrotemporally modulated ‘ripple’ sounds are considered the auditory analogs of ‘Gabor’ patches, which have been extensively used to study rule-based learning (Visscher, Kaplan, Kahana, & Sekuler, 2007). 2) ‘Ripple’ sounds have been extensively studied in animals and humans using electrophysiological (single neuron), neuroimaging, and behavioral methods (Depireux, Simon, Klein, & Shamma, 2001; Langers, Backes, & van Dijk, 2003; Schönwiesner & Zatorre, 2009). These studies reveal that the representation of each dimension is independent of the other (neurally separable) at multiple levels of auditory processing (Depireux et al., 2001; Langers et al., 2003; Schönwiesner & Zatorre, 2009). 3) Low-frequency spectral and temporal modulations are inherent to natural sounds (Santoro et al., 2014; Singh & Theunissen, 2003); therefore, the ripple sounds are considered building blocks of complex auditory signals that are relevant to the human auditory soundscape. 4) A previous study found that short-term memory processes are comparable across visual and auditory modalities when Gabor patches and auditory ripple sounds were used as stimuli (Visscher et al., 2007). It is important to note that relative to vision, the auditory dimensions are more difficult to verbalize (especially the spectral modulation dimension). It is important to note that the presence of a ‘verbal system’ isn’t a prerequisite for rule-based learning. As an example, macaques can learn rule-based category structures, suggesting that a lack of a language system does not preclude rule-based learning (Smith et al., 2012).

Computational modeling of perceptual and decisional processes—To elaborate on strategy differences in auditory rule-based learning related to age, we examine decisional and perceptual processes using multidimensional (MD) signal detection theory (Ashby & Townsend, 1986; Maddox & Ashby, 1993). In this framework, each presentation of a single stimulus yields a distinct perceptual effect that, across trials, is represented by a multivariate normal distribution of perceptual effects (Ashby & Townsend, 1986). Changes in the perceptual variances are associated with perceptual selectivity with smaller variances being associated with more veridical perception. In this framework, we can also examine decisional processes that involve constructing decision bounds (defined in detail later) that carve up the perceptual space into separate response regions. Perceptual and decisional processes are theoretically independent, and therefore have unique, identifiable parameters (Ashby & Townsend, 1986; Green & Swets, 1966; Maddox & Ashby, 1993).

Auditory perceptual development has been relatively well-studied in human and animal models (Sanes & Woolley, 2011). Sanes and Woolley (2011) summarize two major themes that emerge from an extensive survey of the literature on auditory perceptual maturation: a) perceptual development continues through adolescence and b) adult-like performance varies based on the particular auditory feature/dimension examined. For example, frequency modulation detection is adult-like before adolescence, but amplitude modulation detection exhibits a more protracted development (Banai, Sabin, & Wright, 2011). Perceptual learning of even basic auditory features is not adult-like by adolescence (Huyck & Wright, 2011).

Given these findings, it is important to assess the relative contribution of emerging perceptual abilities and decisional processes to individual differences in auditory rule-based category learning. To this end, in the current study, we examine changes in perceptual processing and decisional processes (strategy use) as a function of age.

Hypotheses—To date, no studies have examined the development of rule-based category learning in the auditory domain, and no studies in vision or audition have examined rule-based category learning in adolescents. These critical gaps in our understanding of rule-based category learning motivate the current study. We predict that adults will *outperform* children and adolescents in the rule-based auditory learning task; computational models will corroborate the accuracy results by indicating more optimal, multidimensional rule use and more veridical percept in adults relative to adolescents and children. The bases of these predictions are that the PFC, the critical component of rule-based processing is not fully mature until adulthood. Additionally, we hypothesize that individual difference in executive function, as measured by performance on the Wisconsin Card Sorting Test (WCST) will predict learning accuracy as well as optimal strategy use.

Methodology

Ethics Statement

The current study was approved by the Institutional Review Board at the University of Texas at Austin. Written informed consent was obtained from all participants. For children and adolescents under 18 years of age, written informed consent was obtained from parents prior to participation.

Participants

All study participants had 6 years of past music training and were not currently practicing a musical instrument. This is due to the fact that previous studies have shown that music experience influences auditory learning (Kraus & Chandrasekaran, 2010; Skoe & Kraus, 2012; Strait, Kraus, Parbery-Clark, & Ashley, 2010). Participants had normal hearing confirmed via audiometric screening test 20 dB in both ears for octaves ranging from 1000 to 4000 Hz.

Forty-one children and adolescents with normal hearing were recruited from St. Elias Orthodox Church School in Austin, TX. Nineteen adults with normal hearing were recruited from the University of Texas at Austin. All participants reported no diagnosis of a developmental disability and had above average nonverbal intelligence, which was confirmed by the Kaufman Brief Intelligence Test, Second Edition (KBIT-2; Kaufman & Kaufman, 2004). All standard scores for KBIT-2 have a mean of 100 and standard deviation of 15, with the average range being defined as scores between 85 and 115. All participants had average or above average verbal and nonverbal memory, which was confirmed by the Test of Memory and Learning, Second Edition (TOMAL-2; Reynolds & Voress, 2007). The TOMAL-2 is another standardized assessment tool with a mean of 100 and standard deviation of 15, with normal limits ranging from 85–115. The KBIT-2 and TOMAL-2 were used as screening tools (see Table 1 for scores). The WCST was used as an independent

measure of individual differences in executive function. Upon completion of all experimental procedures, children received \$10 compensation per hour as well as a prize for their participation. Adult participants were compensated \$10 per hour for their participation.

Background Measures

General Background Questionnaire—Demographic information was collected from the adult participants and the child participants via parents, in order to control for socioeconomic status and to confirm that no individual had a developmental disability.

Yale Journal of Sociology Four Factor Index of Social Status—The Yale Journal of Sociology Four Factor Index of Social Status (Adams & Weakliem, 2011) was utilized to calculate reliable socioeconomic scores for each participant and control for socioeconomic environment. The Social Stratum for each participant was derived by a four factor index of social status which equals: occupation $\times$ education $\times$ gender $\times$ marital status. All participants' family social strata in this study fell into two categories: 1) medium business, minor professional, technical (Social Stratum range=54-40) or 2) major business and professional (Social Stratum range=66-55).

Kaufman Brief Intelligence Test-Second Edition (KBIT-2; Kaufman & Kaufman, 2004)—The nonverbal matrices subtest of the KBIT-2 was administered to assess the nonverbal intelligence for all participants (Kaufman & Kaufman, 2004). This assessment tool has been normed for age range= 4.0–90.0. The standard procedure as described in the administrator's manual was utilized for testing and scoring.

Test of Memory and Learning, Second Edition (TOMAL-2; Reynolds & Voress, 2007)—This assessment tool has been normed for age range= 5.0–59.11. The test is comprised of 8 core subtests: Memory for Stories, Facial Memory, Word Selective Reminding, Abstract Visual Memory, Object Recall, Visual Sequential Memory, Paired Recall, and Memory for Location. These subtests yield verbal, nonverbal and composite memory indices.

Experiment Materials

Auditory Ripple Stimulus space—All auditory stimuli varied along the two well-defined dimensions of temporal modulation and spectral modulation. A total of 600 stimuli were generated in the spectrotemporal modulation space with 150 being sampled from each of four categories (see Figure 1A). Each category was bivariate normally distributed with the following means and variances and with all covariance terms equal to zero: ($X_1 \sim N(0.33, 0.1)$, $Y_1 \sim N(0.33, 0.1)$); ($X_2 \sim N(0.33, 0.1)$, $Y_2 \sim N(0.75, 0.1)$); ($X_3 \sim N(0.75, 0.1)$, $Y_3 \sim N(0.33, 0.1)$); and ($X_4 \sim N(0.75, 0.1)$, $Y_4 \sim N(0.75, 0.1)$). The coordinates in the abstract two-dimensional space were then transformed linearly onto an acoustic space defined by temporal modulation frequency dimension ([4, 12 Hz]) and spectral modulation frequency dimension with bounds of ([0.1, 2 cyc/oct]) using Equation 1:

$$S_{acoustic} = S_{abstract} \begin{bmatrix} 8 \text{ Hz} & 0 \\ 0 & 1.9 \text{ cyc/oct} \end{bmatrix} + \begin{bmatrix} 4 \text{ Hz} & 0.1 \text{ cyc/oct} \\ \vdots & \vdots \\ 4 \text{ Hz} & 0.1 \text{ cyc/oct} \end{bmatrix}$$

Where $S_{abstract}$ was a 150×2 matrix containing the original abstract coordinates.

These frequency bounds were selected for their correspondence with the spectrotemporal properties of naturally-occurring complex sounds (Schönwiesner & Zatorre, 2009; Woolley, Fremouw, Hsu, & Theunissen, 2005). The frequency ranges for the auditory stimuli were specifically chosen to reflect the low-pass filter characteristics of the auditory system (Schönwiesner & Zatorre, 2009; Woolley et al., 2005). The spectrotemporally modulated sounds were generated using a custom MATLAB script with the following parameters: duration = 1 s; phase = 0° ; F_0 = 200 Hz; spectral bandwidth = 3 oct; amplitude modulation depth = 30 dB; RMS amplitude = 82 dB; sampling rate = 44.1 kHz.

Because the category means form a rectangle in the spectrotemporal modulation space and the variance is fixed across categories along the spectral modulation and temporal modulation dimensions, the optimal strategy is to set a criterion along the spectral modulation dimension, and to set a criterion along the temporal modulation dimension. The optimal spectral modulation decision criterion is 1.05 cyc/oct, and the optimal temporal modulation criterion is 8 Hz. These are then combined in such a way that low spectral and low temporal modulation is associated with category A, high spectral and low temporal modulation is associated with category B, low spectral and high temporal modulation is associated with category C, and high spectral and high temporal modulation is associated with category D.

Wisconsin Card Sorting Task (WCST)—The Wisconsin Card Sorting Test (WCST) is a gold standard measure of executive function in neuropsychology (Heaton, 1993) and examines single-dimension rule learning and set shifting by challenging participants to sort cards by number, color, and shape. Neuroimaging work has demonstrated that implementation of the Wisconsin Card Sorting Test (WCST) increases demand on attention, working memory and cognitive control, and in turn, activates the right DLPFC and ACC (Lie, Specht, Marshall, & Fink, 2006). The DLPFC and the ACC are integral components of the rule-based learning system (Adams & Weakliem, 2011). Therefore, the WCST is a valid index of individual differences in executive function, specific to those executive functions (e.g. attention, working memory, cognitive control) engaged in rule-based processing.

Procedure

Before the categorization experiment was completed, all participants signed an informed consent document and completed an audiometric screening test following a protocol approved by the University of Texas at Austin Institutional Review Board. Adult participants, and parents of child and adolescent participants under 18 years of age, completed background forms confirming that they had no developmental disability. The experiment took place in a sound-attenuated booth using E-Prime 2.0 software (Schneider,

Eschman, & Zuccolotto, 2002). The ripple stimuli were bilaterally presented to participants through Sennheiser headphones at a fixed, comfortable loudness level of ~ 70 dB SPL, as measured by a sound level meter. This level was maintained across all participants.

Session 1: Category Learning & Intelligence Measures—All participants were instructed to listen to sounds in each trial and categorize them into one of the four categories using number keys 1, 2, 3, and 4. Following each trial-by-trial response, feedback was presented immediately: “Right,” or “Wrong” with an alien who was shown smiling or frowning, respectively. Each stimulus was presented once following a randomized sequence for individual sessions, resulting in 600 trials for each learning session. In order to maintain the attention of the subjects, the experiment had a solar system theme. All subjects were told they would be playing a solar system exploration game (see Figure 1b for an example of the task). Following the rule-based auditory categorization task, the subjects were given a break before completing the KBIT-2. The first experimental session lasted approximately one hour.

Session 2: Memory and Executive Function Measures—Approximately 1 week after the completion of the ripple categorization task and the KBIT-2, the subjects returned to complete the TOMAL-2 and the WCST. The second experimental session lasted approximately one hour and a half.

Results

Auditory Categorization Performance Analyses

A mixed effects regression analysis was conducted to predict the accuracy of producing a correct response using the lmer program with binomial logit link (Bates, Maechler, & Bolker, 2012). In mixed effects modeling analyses fixed-effects parameters are a set of possible levels of a covariate that are fixed and reproducible, while random effects represent a random sample from the set of all possible levels of a variable. We have employed similar analyses methods in previous papers involving auditory category learning (Chandrasekaran et al., 2015; Chandrasekaran, Yi, et al., 2014; Yi et al., 2014).

In the current analysis, trial-by-trial accuracy was set as the dependent variable. For each participant, the response to each trial was coded as “correct” or “incorrect”. The log odds of producing a correct response in a given trial were estimated for fixed effects of age group (child, adolescent, and adult; where the adolescent group was the reference level), trial (increasing from 1 to 600; mean centered to 0), and their interaction terms. In order to correct for individual variability in learning trajectory, by-participant random slopes for trial number were also included. Four simple effects were estimated by the analysis. First, the intercept modeled the log odds of a correct response for adolescent group categorization accuracy in a given trial. Second, the adult effect modeled the difference between the adult group relative to the adolescent group categorization accuracy. Third, the child effect modeled the difference between the child group relative to the adolescent group categorization accuracy. Fourth, the trial effect modeled the adolescent group categorization accuracy across learning trials. The adult group by trial interaction effect allowed for comparison between the adult and adolescent groups, across learning trials. Similarly, the

child group by trial interaction effect allowed for comparison between the child group and the adolescent group across learning trials. An estimate of zero for the interaction effect would suggest that the difference between two simple effects was fully accounted for by the separate effects. A significantly non-zero estimate would suggest the degree to which the two simple effects fail to account for the difference.

Effect of age on auditory categorization performance

Figure 2a displays the learning curves for children, adolescents and adults. As observed in Table 2, the adult group effect was significant ($p < .001$), indicating that adults had higher auditory categorization accuracy relative to adolescents. The child group effect was approaching significance ($p = 0.08$), indicating that auditory categorization accuracy in the child group was marginally significantly lower than the adolescent group accuracy. The trial effect was significant ($p < .001$), suggesting that the probability of an accurate response increased as a function of trial for adolescents. The adult group $\times$ trial interaction was significant ($p = .002$), suggesting that the accuracy increase as a function of trials was greater for adults relative to adolescents. The child group $\times$ trial interaction was significant ($p < .001$), suggesting that the accuracy increase over learning trials was smaller for children relative to adolescents¹.

Model-Based Analyses

The accuracy-based analyses suggest that young adults show faster auditory rule-based learning relative to adolescents, whereas children show slower auditory rule-based learning relative to adolescents. Although informative in their own right, and often the only performance measure examined, accuracy rates do not provide information about the types of strategies that participants use to solve the task. To provide insights into the critical decisional processes involved in auditory category learning, we utilized computational models of the task that allow us to characterize the nature of the response strategy an individual is applying in a given task. We apply a series of decision bound models originally developed for application in the visual domain (Ashby & Maddox, 1993; Maddox & Ashby, 1993) and recently extended to the auditory domain (Chandrasekaran, Yi, et al., 2014; Maddox & Chandrasekaran, 2014; Maddox et al., 2013) on a block-by-block basis at the individual participant level because of problems with interpreting fits to aggregate data (Ashby, Maddox, & Lee, 1994; Estes, 1956; Maddox, 1999). We assume that each point in the two-dimensional spectrotemporal modulation space (displayed in Figure 1a) accurately describes the average (or mean) perceptual effects for each stimulus. We estimate parameters associated with the decision strategy being utilized by the participant. Each model assumes that decision-bound (or category boundaries created by the participant as they learn the categories) were used to classify stimuli into each of the four sound categories.

Here we provide a brief description of the decision processing assumptions of each model. Details are available in numerous previous publications (Ashby & Maddox, 1993; Maddox

¹Although the accuracy rates for children were low, they were significantly above chance in all 6 blocks (all p values $< .001$). For completeness, we also compared block-by-block performance for adolescents and adults against chance and in all six blocks for each age group the effects were significant (all p values $< .001$).

& Chandrasekaran, 2014; Maddox & Ashby, 1993). We applied four classes of models. The first two assume that the participant uses a multidimensional strategy. The conjunctive (CJ), rule-based model assumes that the participant sets criteria along the temporal modulation dimension and along the spectral modulation dimension and that these are then combined to determine category membership. This model contains two free decision criterion parameters and a single free parameter that represents noise in the perceptual system. The striatal pattern classifier (SPC) is a model that assumes a reflexive multidimensional strategy. This model captures many of the characteristics of implicit, procedural-based processing. However, at the computational level this is a prototype model in which each category is represented by a single prototype in the temporal-spectral modulation space and each exemplar is classified into the category with the most similar prototype. Because the location of one of the prototypes can be fixed, and since a uniform expansion or contraction of the space will not affect the location of the resulting response region partitions, the multidimensional reflexive (prototype) model contains five free decision parameters that determine the location of the prototypes, and a single free parameter that represents noise in the perceptual system. Unidimensional (UD) models assume that the participant sets criteria along the temporal modulation or spectral modulation dimension that are then used to determine category membership. For example, the Unidimensional Temporal modulation model (UDX) assumes that the participant sets three criteria along the temporal modulation dimension that are used to separate the stimuli into those that are of low, medium-low, medium-high, or high temporal modulation. Importantly, this model ignores the spectral modulation dimension. Although a large number of versions of this model are possible, we explored the four variants of the model that made the most reasonable assumptions regarding the assignment of category labels to the four response regions. Using the convention that the first, second, third and fourth category labels are associated with low, medium-low, medium-high and high temporal modulation respectively, the 4 variants were: 1234, 1243, 2134, and 2143. The Unidimensional Spectral model (UDY) assumes that the participant sets three criteria along the spectral modulation dimension. The model assumes that the three criteria along the spectral modulation dimension are used to separate the stimuli into those that are of low, medium-low, medium-high, or high spectral modulation. Importantly, this model ignores the temporal modulation dimension. Although a large number of versions of this model are possible, we explored the four variants of the model that made the most reasonable assumptions regarding the assignment of category labels to the four response regions. Using the convention that the first, second, third and fourth category labels are associated with low, medium-low, medium-high and high spectral modulation, respectively, the 4 variants were: 1324, 1342, 3124 and 3142. These models contains three free decision criterion parameters and a single free parameter that represents noise in the perceptual system.

Notice that the UDX and UDY models assume that the participant stores three separate decision criteria in memory and uses these to generate categorization responses. This places an increased load on working memory and may be unlikely in children, adolescents and even some younger adults. In light of this fact, we explored two other classes of unidimensional models. The first class assumes that the participant sets *one criterion* along the temporal modulation dimension. This criterion separates the space into stimuli “low” vs.

“high” in temporal modulation. Four versions of this model were applied to the data: 1) Low temporal modulation was associated with category 1, and high temporal modulation led to guessing between categories 2, 3, and 4 with the guessing probabilities being free parameters, 2) Low temporal modulation was associated with category 2 and high temporal modulation led to guessing between categories 1, 3 and 4. 3) High temporal modulation was associated with category 3, and low temporal modulation led to guessing between categories 1, 2, and 4. 4) High temporal modulation was associated with category 4, and low temporal modulation led to guessing between categories 1, 2, and 3. The second class assumes that the participant sets one criterion along the Spectral modulation dimension. This criterion separates the space into stimuli “low” vs. “high” in spectral modulation. Four versions of this model were applied to the data. 1) Low spectral modulation was associated with category 1, and high spectral modulation led to guessing between categories 2, 3, and 4. 2) Low spectral modulation was associated with category 3 and high temporal modulation led to guessing between categories 1, 2 and 4. 3) High spectral modulation was associated with category 2, and low spectral modulation led to guessing between categories 1, 3, and 4. 4) High spectral modulation was associated with category 4, and low spectral modulation led to guessing between categories 1, 2, and 3. All 8 of these models include 1 decision criterion parameter, 2 guessing parameters (the third was fixed since the three guessing probabilities must sum to 1.0), and one perceptual noise parameter for a total of 4 parameters.

The fourth model is a random responder model that assumes that the participant guesses on each trial. All models except the random responder model also included a perceptual noise parameter that provided an estimate of trial-by-trial perceptual fluctuations.

Each of the models (except the random responder model) included a perceptual noise parameter. Figure 1 displays a scatterplot of the spectral and temporal modulation values for each stimulus. In the models we assumed that these values denote the mean perceptual effects of the stimuli with variability in the trial-by-trial perceptual effects being estimated from the data. We assume that the perceptual variance along spectral modulation is identical to the perceptual variance along temporal modulation, which this value is identical across all stimuli, and that perceptual variability along the two dimensions is uncorrelated. Critically, a smaller perceptual variance is associated with a more veridical percept, and the decision processing assumptions associated with the different models are independent of one another (Green & Swets, 1966; Maddox & Ashby, 1993). The fact that we can separately estimate perceptual from decisional processes is critical as it allows us to examine the developmental trajectory of each separately.

The models were fit to the category learning data from each trial on a block-by-block basis by maximizing negative log-likelihood and the best fitting model was identified by comparing BIC values for each model (Akaike, 1974; Schwarz, 1978). BIC penalizes models with more free parameters. For each model, i , BIC is defined as:

$$BIC_i = -2\ln L_i + V_i \ln(n)$$

where n is the number of trials being fit, L_i is the maximum likelihood for model i , and V_i is the number of free parameters in the model. Smaller BIC values indicate a better fit to the data².

Results of the Model-Based Analyses—To assess the effect of age on computational strategies, we employed mixed effects regression analyses. Because the optimal strategy is conjunctive, for each participant, conjunctive strategy use during each block was coded as “TRUE” or “FALSE”, depending upon whether the conjunctive strategy provided the best fit to the data in each block. The proportion of conjunctive strategy use was then set as the dependent variable. The log odds of using a conjunctive strategy in a given block was estimated for fixed effects of age group (child, adolescent, and adult; where the adolescent group was the reference level), block (increasing from 1 to 6; mean centered to 0) and their interaction terms. In order to correct for individual variability in conjunctive strategy use, by-participant random slopes for block number were included. Four simple effects were estimated by the analysis. First, the intercept modeled the log odds of conjunctive strategy use for the adolescent group in a given block. Second, the child effect modeled the difference between the child group and the adolescent group conjunctive strategy use. Third, the adult effect modeled the difference between the child group and the adolescent group conjunctive strategy use. Fourth, the block effect modeled conjunctive strategy use across learning blocks for the adolescent group. The adult group by block interaction effect allowed for comparison between the adult group and the adolescent group across learning blocks. The child group by block interaction effect allowed for comparison between the child group and the adolescent group across learning blocks.

Effect of age on conjunctive strategy use during auditory category learning

The proportion of participants in each age group best fit by each model type is displayed in Figure 2b, and the mixed modeling results are displayed in Table 3. The block effect was significant ($p < .001$), suggesting that conjunctive strategy use increases as a function of block for adolescents. The adult group $\times$ block interaction was significant ($p < .001$), suggesting that the increase in conjunctive strategy use over blocks was greater for adults than adolescents. The child group $\times$ block interaction was significant ($p < .001$), suggesting that the increase in conjunctive strategy use over blocks was less for the child group, relative to the adolescent group.

Effect of age on perceptual processing during auditory category learning

As outlined in the model description section above, a strength of our modeling approach is that we obtain separate estimates of perceptual processes. This allows us to examine the effects of development on decisional and perceptual processes separately. To examine the effects of age (children, adolescents, and young adults) on perceptual processing, we conducted a linear mixed effects regression analysis (see Table 4). In this analysis,

²We also determined the best fitting model from Akaike’s Information Criterion (AIC) that is also derived from maximum likelihood estimation. In general, AIC tends to be more liberal with models with more free parameters often providing the best account of the data. Because of the very large number of models being examined in this manuscript, we decided to be more conservative with a bias toward models with fewer free parameters. Importantly though, all of the qualitative patterns observed with BIC were mirrored with AIC.

perceptual noise was set as the dependent variable. The log odds of perceptual noise were estimated for fixed effects of age group (child, adolescent, and adult; where the adolescent group was the reference level) and block (increasing from 1 to 6; mean centered to 0), and their interaction terms. In order to correct for individual variability in learning trajectory, by-participant random slopes for trial number were included.

The adult group effect was significant ($p < .001$), indicating that perceptual noise was less for the adult group, relative to adolescent group. The child group effect was significant ($p < .001$), suggesting that overall perceptual noise was greater for the child group, relative to the adolescent group. The block effect was not significant ($p = 0.84$), indicating that perceptual noise did not increase as a function of block for the adolescent group. There were no significant two-way interactions, suggesting that perceptual noise did not increase significantly more or less for the adult group and the child group across learning blocks.

In addition, we estimated whether final block accuracy is related to optimal strategy use as well as perceptual variance (in the same model; see Table 5). In this analysis, final block accuracy was set as the dependent variable. The log odds of producing a correct response in a given trial of the final block were estimated for fixed effects of age group (child, adolescent, and adult; where the adolescent group was the reference level), final block strategy (conjunctive strategy vs. non-conjunctive strategy; where conjunctive strategy was the reference level), final block perceptual noise, and their interaction terms. In order to correct for individual variability in learning trajectory, by-participant random slopes for trial number were included.

The adult group effect was not significant ($p = 0.32$), indicating that final block accuracy for adult conjunctive strategy users did not differ from the final block accuracy of adolescent conjunctive strategy users. The child group effect was not significant ($p = 0.79$), suggesting that final block accuracy for child conjunctive strategy users did not differ from final block accuracy of adolescent conjunctive strategy users. The final block, non-conjunctive strategy effect was significant ($p < .001$), suggesting that non-conjunctive strategy use negatively predicted final block accuracy. The final block noise effect was also significant ($p < .001$), suggesting that perceptual noise negatively predicted final block accuracy. There were no significant two or three way interactions, suggesting that perceptual noise and non-conjunctive strategy use negatively predicts final block accuracy for all groups to a similar extent.

Individual Differences in Executive Function

WCST as a Predictor of Auditory Category Learning Success and Conjunctive Strategy Use

Mixed effects regression analysis was conducted to determine if overall WCST performance was predictive of participant's overall accuracy using the lmer program with binomial logit link (see Table 6). For each participant, the response to each trial was coded as "correct" or "incorrect". The log odds of producing a correct response in a given trial were estimated for fixed effects of trial (increasing from 1 to 600; mean centered to 0), overall accuracy in the Wisconsin Card Sorting Test, and their interaction terms. In order to correct for individual variability in executive flexibility, by-participant, as well as by-age random slopes for block number were included. Age was treated as a continuous variable.

Three simple effects were estimated by the analysis. First, the intercept modeled the log odds of correct response in a given auditory category learning trial. Second, trial effect estimated the extent to which increase in trials predicted category learning accuracy. Third, the WCST effect estimated the extent to which WCST accuracy predicted auditory category learning accuracy. The trial by WCST two-way interaction indicated the extent to which overall WCST was predictive of higher rule-based auditory category learning across trials.

The trial by WCST two-way interaction was significant ($p < .001$), indicating that the interaction of trials and overall WCST performance was predictive of higher rule-based auditory learning accuracy. This suggests better executive flexibility results in more successful rule-based auditory category learning. The overall WCST performance effect was significant ($p = .04$), suggesting that higher overall WCST performance is significantly predictive of higher rule-based auditory categorization accuracy. The trial effect was significant ($p < .001$), suggesting that probability of an accurate response increases as a function of trial.

A second mixed effects regression analysis was conducted to determine whether WCST performance was predictive of participant's proportion of conjunctive strategy use, using the lmer program with binomial logit link (see Table 7). Here, the dependent variable was set as conjunctive strategy use. The log odds of using a conjunctive strategy in a given block were estimated for fixed effects of block (increasing from 1 to 6; mean centered to 0), overall Wisconsin Card Sorting Task accuracy, and their interaction terms. In order to correct for individual variability in executive flexibility, by-participant, as well as by-age random slopes for block number were included. Age was treated as a continuous variable. Three simple effects were estimated by the analysis. First, the intercept modeled the log odds of conjunctive strategy use for a given block. Second, block effect estimated the extent to which increase in blocks predicted conjunctive strategy use. Third, the WCST effect estimated the extent to which WCST accuracy predicted conjunctive strategy use. The block by WCST two-way interaction estimated the extent to which overall WCST was predictive of conjunctive strategy use across blocks.

The block by overall WCST accuracy two-way interaction was significant ($p < .001$), indicating that the interaction of learning and overall WCST accuracy was predictive of greater CJ strategy use. The WCST effect was significant ($p < .001$), suggesting that higher overall WCST accuracy is significantly predictive of more conjunctive strategy use. The block effect was significant ($p < .001$), suggesting that probability of using a conjunctive strategy increases as a function of block.

Discussion

We examined the development of rule-based category learning in the auditory domain using behavioral and computational modeling analyses. We additionally examined the extent to which variation in executive function predicted individual differences in accuracy and computational strategies used during rule-based auditory category learning. As hypothesized, the results show that rule-based auditory category learning improves with age, with young adults outperforming children and adolescents, and adolescents outperforming

children. Computational modeling analyses indicated that the use of the task-optimal strategy (i.e. conjunctive strategy) and veridical perception (i.e., lower perceptual noise) increased with age. Irrespective of age, the use of task-optimal strategy and lower perceptual noise related to enhanced learning accuracy in the final block of training. Finally, individual differences in executive function significantly predicted higher overall rule-based auditory categorization accuracy, as well as use of the optimal conjunctive strategy. We discuss each of these findings in detail.

Development of Auditory Category Learning and Rule-based Strategies

The accuracy results show a protracted development of rule-based auditory categorization in the auditory domain. Adults outperform adolescents and children; adolescents outperform children. Previous work on auditory perceptual learning has yielded a similar protracted developmental profile (Huyck & Wright, 2011; Sanes & Woolley, 2011). Our computational modeling results show that the perceptual variance associated with the auditory ripple stimuli reduces with age, indicating more veridical perception. Computational modeling results indicate an increase in the ability to apply optimal conjunctive rules with age. Children are the least likely to use conjunctive rule-based strategies, instead focusing more on unidimensional strategies, whereas young adults are the most likely to use conjunctive strategies. Adolescents show intermediate levels of conjunctive strategy use. The protracted development of rule-based categorization and sub-optimal strategy use, demonstrated in the current study, is consistent with developmental evidence from the visual domain (Huang-Pollock et al., 2011; Minda et al., 2008; Rabi et al., 2015; Rabi & Minda, 2014). Our results suggest that poor rule-based learning is not modality-specific and may reflect slow maturation of a common rule-based network (i.e. the DLPFC and ACC). Previous studies in the visual domain have not examined adolescent performance on rule-based tasks. Our results unequivocally demonstrate that ability to learn complex rule-based strategies is not complete in adolescence. To that end, these findings may reflect a more general slowly maturing prefrontal network that is critical to hypothesis-testing across sensory modalities.

Multiple Learning Systems in Audition

As discussed in the introduction section, numerous studies have examined the development of auditory category structures that are optimally learned by the implicit learning system. A general finding is a reduction in learning ability as a function of age. Here we show that on an auditory learning task that we posit is optimally learned via a slow maturing, explicit, rule-based learning system, there is an increase in learning ability as a function of age. The present results also reveal that executive function is associated with successful auditory rule-based strategy use, and ultimately auditory rule-based category learning success.

This study represents the first to examine the developmental trajectory of auditory rule-based category learning. Our results suggest that in addition to implicit learning processes, human auditory learning is also mediated by an explicit, hypothesis-testing system. These results are consistent with predictions made by dual-learning system approaches in vision that posit that implicit and explicit learning systems, with the dominant system determined by the nature of the underlying category structure to be learned (Ashby et al., 2011). Dual-learning system models assume that the two systems compete throughout learning, but that

initial learning is dominated by the explicit, hypothesis-testing system (Huang-Pollock et al., 2011).

Because there is an initial bias toward hypothesis-testing strategies as a function of age, it makes sense to begin our examination of the development of auditory category learning with rule-based categories. Even so, a more complete understanding of the development of auditory category learning will require further examination of information-integration category structures. A few developmental studies have been conducted in the visual domain. These studies reveal that both rule-based and information-integration category learning improves with age (Huang-Pollock et al., 2011; Rabi et al., 2015; Visser & Raijmakers, 2012). Children's poorer learning of visual rule-based categories has been explained by a lack of ability to identify successful rules, and inhibit unsuccessful rules, while children's poorer learning of visual information-integration categories has been explained by a lack of ability to *disengage* from the hypothesis-testing system. Future work is needed to examine the developmental trajectory of auditory information-integration category learning and to compare it with auditory rule-based category across development.

The current study is the first to establish the developmental profile of rule-based categorization in audition. In addition, by studying adolescent population we extend prior work in vision that has compared rule-based category learning in children and adults. We find that optimal learning in the rule-based auditory category learning task necessitates the use of conjunctive strategy, which is dependent on executive function and age. These results suggest that rule-based category learning is not modality specific. Our computational modeling provides independent estimations of perceptual and decisional processes, both of which relate to rule-based learning success. Given prior findings of poorer executive function in children with language and learning disorders (Hutchinson, Bavin, Efron, & Sciberras, 2012; Petrucci, Bavin, & Bretherton, 2012; Singer & Bashir, 1999), as well as perceptual impairments (Wright et al., 1997), further examination of performance on rule-based auditory category learning may provide a mechanistic and computational basis for understanding the mechanisms underlying language and learning disorders.

Conclusion

The current findings demonstrate the protracted development of auditory categorization performance from childhood through early adulthood. Our results further highlight the role that perceptual and prefrontally modulated cognitive processes play in successful auditory category learning. Prefrontally modulated neurocognitive processes, such as executive function, are necessary in order to fully utilize and engage hypothesis-testing during rule-based auditory category learning.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this paper was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under Award Number R01DC013315 (BC). The content is solely the responsibility of the authors and does not represent the official view of the National Institutes of Health. The authors thank the children, parents, university students, and professors who made this study possible. The authors would also like to thank Han-Gyol Yi, Zilong Xie, Rachel Tessmer, Nicole Tsao, and other members of the SoundBrain Laboratory for significant contributions in stimuli development, data collection and processing.

References

- Adams J, Weakliem DL. B. Hollingshead's "Four Factor Index of Social Status": From Unpublished Paper to Citation Classic. *Yale Journal of Sociology*. 2011 Aug.:11.
- Akaike H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*. 1974; 19(6):716–723.
- Ashby F, Paul E, Maddox W. COVIS. Formal approaches in categorization. 2011:65–87.
- Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM. A neuropsychological theory of multiple systems in category learning. *Psychological review*. 1998; 105(3):442. [PubMed: 9697427]
- Ashby FG, Maddox WT. Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*. 1993; 37(3):372–400.
- Ashby FG, Maddox WT. Human category learning 2.0. *Annals of the New York Academy of Sciences*. 2011; 1224(1):147–161. [PubMed: 21182535]
- Ashby FG, Maddox WT, Lee WW. On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science*. 1994; 5(3):144–151.
- Ashby FG, Townsend JT. Varieties of perceptual independence. *Psychological review*. 1986; 93(2):154. [PubMed: 3714926]
- Banai K, Sabin AT, Wright BA. Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. *Hearing research*. 2011; 280(1):219–227. [PubMed: 21664958]
- Bates D, Maechler M, Bolker B. lme4: Linear mixed-effects models using Eigen and Eigen. 2012
- Bradlow AR, Pisoni DB, Akahane-Yamada R, Tohkura Yi. Training Japanese listeners to identify English /r/ and /l/. IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*. 1997; 101(4):2299–2310. [PubMed: 9104031]
- Casey B, Tottenham N, Liston C, Durston S. Imaging the developing brain: what have we learned about cognitive development? *Trends in cognitive sciences*. 2005; 9(3):104–110. [PubMed: 15737818]
- Chandrasekaran B, Koslov SR, Maddox WT. Toward a dual-learning systems model of speech category learning. *Frontiers in Psychology*. 2014; 5
- Chandrasekaran B, Yi HG, Blanco NJ, McGeary JE, Maddox WT. Enhanced Procedural Learning of Speech Sound Categories in a Genetic Variant of FOXP2. *The Journal of Neuroscience*. 2015; 35(20):7808–7812. [PubMed: 25995468]
- Chandrasekaran B, Yi HG, Maddox WT. Dual-learning systems during speech category learning. *Psychonomic Bulletin & Review*. 2014; 21(2):488–495. [PubMed: 24002965]
- Crone EA, Donohue SE, Honomichl R, Wendelken C, Bunge SA. Brain regions mediating flexible rule use during development. *The Journal of Neuroscience*. 2006; 26(43):11239–11247. [PubMed: 17065463]
- Depireux DA, Simon JZ, Klein DJ, Shamma SA. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *Journal of neurophysiology*. 2001; 85(3):1220–1234. [PubMed: 11247991]
- Estes WK. The problem of inference from curves based on group data. *Psychological bulletin*. 1956; 53(2):134. [PubMed: 13297917]

- Gabrieli JD, Brewer JB, Desmond JE, Glover GH. Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science*. 1997; 276(5310):264–266. [PubMed: 9092477]
- Gifford AM, Cohen YE, Stocker AA. Characterizing the impact of category uncertainty on human auditory categorization behavior. *PLoS computational biology*. 2014; 10(7):e1003715. [PubMed: 25032683]
- Green, D.; Swets, J. Signal detection theory and psychophysics. New York: 1966.
- Heaton, RK. Wisconsin card sorting test: computer version 2. Odessa: Psychological Assessment Resources; 1993.
- Huang-Pollock CL, Maddox WT, Karalunas SL. Development of implicit and explicit category learning. *Journal of experimental child psychology*. 2011; 109(3):321–335. [PubMed: 21377688]
- Hutchinson E, Bavin E, Efron D, Sciberras E. A comparison of working memory profiles in school-aged children with Specific Language Impairment, Attention Deficit/Hyperactivity Disorder, Comorbid SLI and ADHD and their typically developing peers. *Child Neuropsychology*. 2012; 18(2):190–207. [PubMed: 21919558]
- Huyck JJ, Wright BA. Late maturation of auditory perceptual learning. *Developmental science*. 2011; 14(3):614–621. [PubMed: 21477199]
- Kaufman, AS.; Kaufman, NL. Kaufman brief intelligence test. 2. 2004.
- Kraus N, Chandrasekaran B. Music training for the development of auditory skills. *Nature reviews neuroscience*. 2010; 11(8):599–605. [PubMed: 20648064]
- Kuhl PK. Early language acquisition: cracking the speech code. *Nature reviews neuroscience*. 2004; 5(11):831–843. [PubMed: 15496861]
- Langers DR, Backes WH, van Dijk P. Spectrotemporal features of the auditory cortex: the activation in response to dynamic ripples. *Neuroimage*. 2003; 20(1):265–275. [PubMed: 14527587]
- Lichtenstein D, Goldstein I, Mourgeon E, Cluzel P, Grenier P, Rouby JJ. Comparative diagnostic performances of auscultation, chest radiography, and lung ultrasonography in acute respiratory distress syndrome. *The Journal of the American Society of Anesthesiologists*. 2004; 100(1):9–15–19–15.
- Lie CH, Specht K, Marshall JC, Fink GR. Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Neuroimage*. 2006; 30(3):1038–1049. [PubMed: 16414280]
- Lively SE, Logan JS, Pisoni DB. Training Japanese listeners to identify English/r/and/l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*. 1993; 94(3):1242–1255. [PubMed: 8408964]
- Maddox W, Chandrasekaran B. Tests of a dual-system model of speech category learning. *Bilingualism: Language and cognition*. 2014; 17(04):709–728.
- Maddox WT. On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception & psychophysics*. 1999; 61(2):354–374. [PubMed: 10089766]
- Maddox WT, Ashby FG. Comparing decision bound and exemplar models of categorization. *Perception & psychophysics*. 1993; 53(1):49–70. [PubMed: 8433906]
- Maddox WT, Chandrasekaran B, Smayda K, Yi HG. Dual systems of speech category learning across the lifespan. *Psychology and aging*. 2013; 28(4):1042. [PubMed: 24364408]
- Maddox WT, Chandrasekaran B, Smayda K, Yi HG, Koslov S, Beevers CG. Elevated depressive symptoms enhance reflexive but not reflective auditory category learning. *cortex*. 2014; 58:186–198. [PubMed: 25041936]
- Maddox WT, David A. 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*. 2005; 31(1):100.
- Maddox WT, Filoteo JV, Lauritzen JS. Within-category discontinuity interacts with verbal rule complexity in perceptual category learning. *Journal of experimental psychology: learning, memory, and cognition*. 2007; 33(1):197.
- Maddox WT, Love BC, Glass BD, Filoteo JV. When more is less: Feedback effects in perceptual category learning. *Cognition*. 2008; 108(2):578–589. [PubMed: 18455155]

- Maddox WT, Molis MR, Diehl RL. Generalizing a neuropsychological model of visual categorization to auditory categorization of vowels. *Perception & psychophysics*. 2002; 64(4):584–597. [PubMed: 12132760]
- Minda JP, Desroches AS, Church BA. Learning rule-described and non-rule-described categories: a comparison of children and adults. *Journal of experimental psychology: learning, memory, and cognition*. 2008; 34(6):1518.
- Nomura E, Maddox W, Filoteo J, Ing A, Gitelman D, Parrish T, Reber P. Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex*. 2007; 17(1):37–43. [PubMed: 16436685]
- Petrucelli N, Bavin EL, Bretherton L. Children with specific language impairment and resolved late talkers: working memory profiles at 5 years. *Journal of Speech, Language, and Hearing Research*. 2012; 55(6):1690–1703.
- Rabi R, Miles SJ, Minda JP. Learning categories via rules and similarity: Comparing adults and children. *Journal of experimental child psychology*. 2015; 131:149–169. [PubMed: 25558860]
- Rabi R, Minda JP. Rule-Based Category Learning in Children: The Role of Age and Executive Functioning. *PLoS one*. 2014; 9(1):e85316. [PubMed: 24489658]
- Reynolds, CR.; Voress, J. *Test of Memory and Learning (TOMAL-2)*. TX: Pro-Ed; 2007.
- Sanes DH, Woolley SM. A behavioral framework to guide research on central auditory development and plasticity. *Neuron*. 2011; 72(6):912–929. [PubMed: 22196328]
- Santoro R, Moerel M, De Martino F, Goebel R, Ugurbil K, Yacoub E, Formisano E. Encoding of natural sounds at multiple spectral and temporal resolutions in the human auditory cortex. *PLoS computational biology*. 2014; 10(1)
- Schacter DL, Wagner AD. Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus*. 1999; 9(1):7–24. [PubMed: 10088896]
- Schneider, W.; Eschman, A.; Zuccolotto, A. *Computer software and manual*. Pittsburgh, PA: Psychology Software Tools Inc; 2002. E-Prime (Version 2.0).
- Schönwiesner M, Zatorre RJ. Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. *Proceedings of the National Academy of Sciences*. 2009; 106(34):14611–14616.
- Schwarz G. Estimating the dimension of a model. *The annals of statistics*. 1978; 6(2):461–464.
- Seger CA. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience & Biobehavioral Reviews*. 2008; 32(2):265–278. [PubMed: 17919725]
- Singer BD, Bashir AS. What are executive functions and self-regulation and what do they have to do with language-learning disorders? *Language, Speech, and Hearing Services in Schools*. 1999; 30(3):265–273.
- Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America*. 2003; 114(6):3394–3411. [PubMed: 14714819]
- Skoe E, Kraus N. A little goes a long way: how the adult brain is shaped by musical training in childhood. *The Journal of Neuroscience*. 2012; 32(34):11507–11510. [PubMed: 22915097]
- Smith JD, Berg ME, Cook RG, Murphy MS, Crossley MJ, Boomer J, Ashby FG. Implicit and explicit categorization: A tale of four species. *Neuroscience & Biobehavioral Reviews*. 2012; 36(10): 2355–2369. [PubMed: 22981878]
- Strait DL, Kraus N, Parbery-Clark A, Ashley R. Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hearing research*. 2010; 261(1):22–29. [PubMed: 20018234]
- Trainor LJ, Trehub SE. A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*. 1992; 18(2):394. [PubMed: 1593226]
- Visscher KM, Kaplan E, Kahana MJ, Sekuler R. Auditory short-term memory behaves like visual short-term memory. *PLoS biology*. 2007; 5(3):e56. [PubMed: 17311472]
- Visser I, Raijmakers ME. Developing representations of compound stimuli. *Frontiers in Psychology*. 2012; 3. [PubMed: 22291676]

- Wong PC, Morgan-Short K, Ettlinger M, Zheng J. Linking neurogenetics and individual differences in language learning: The dopamine hypothesis. *cortex*. 2012; 48(9):1091–1102. [PubMed: 22565204]
- Woolley SM, Fremouw TE, Hsu A, Theunissen FE. Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature neuroscience*. 2005; 8(10):1371–1379. [PubMed: 16136039]
- Wright BA, Lombardino LJ, King WM, Puranik CS, Leonard CM, Merzenich MM. Deficits in auditory temporal and spectral resolution in language-impaired children. 1997
- Yi H-G, Maddox WT, Mumford JA, Chandrasekaran B. The role of corticostriatal systems in speech category learning. *Cerebral Cortex*. 2014 bhu236.

Highlights

- We use computational modeling to examine the development of rule-based auditory category learning.
- We examine the association between rule-based auditory category learning and executive function.
- Auditory category learning accuracy and use of a task-optimal strategy improved with age.
- Executive function coupled with perceptual processes play important roles in successful rule-based auditory category learning.

Figure 1.

(a) Auditory stimulus space. (b) Example of the rule-based auditory category learning task.

Figure 2.

(a) Auditory Categorization Learning Curves. Error bars denote 95% confidence intervals.

(b) Proportion of participants fit by a conjunctive model (CJ), striatal pattern classifier model (SPC), unidimensional model (temporal-x dimension (UDX) or spectral-y dimension (UDY), a one dimension guessing model (1DG), and a random responder model (RAN).

Table 1

Participant demographic information

	Children	Adolescents	Adults
n	20	21	19
age range	7–12	13–19	20–25
age mean	10.20(1.70)	16.4(2.2)	21.4(0.96)
SES	55.83(7.82)	51.05(6.81)	47.61(11.42)
IQ (KBIT-2)	115.60(12.50)	105.14(13.47)	113.89(11.24)
TOMAL-2: verbal	103.80(11.31)	100.57(7.90)	99.32(14.99)
TOMAL-2: nonverbal	110.75(10.88)	109.62(8.87)	111.53(10.21)
TOMAL-2: composite	108.00(10.44)	105.24(7.44)	105.95(13.34)

Note. ‘SES’ is the socio-economic status of the parents as measured by the *General Background Questionnaire*. ‘IQ’ is the standard score of the *Kaufman Brief Intelligence Test, Second Edition (KBIT-2)*. Verbal, nonverbal, and composite memory are all standard scores of the *Test of Memory and Learning, Second Edition (TOMAL-2)*.

Table 2

Results of the mixed effects logistic regression on accuracy as a function of age and trial.

Fixed effects	Estimate	Standard Error	z value	p value
(Intercept)	−0.34	0.09	−3.82	<.001
Adult Group	0.49	0.13	3.70	<.001
Child Group	−0.23	0.13	−1.75	0.08
Trial	.001	.0001	10.09	<.001
Adult Group × Trial	.0004	.0002	3.16	.002
Child Group × Trial	−.0007	.0002	−4.66	<.001

Table 3

Results of the mixed effects logistic regression on conjunctive strategy use as a function of age group and block.

Fixed effects	Estimate	Standard Error	z value	p value
(Intercept)	-4.03	0.91	-4.45	<.001
Adult Group	1.74	1.24	1.40	0.16
Child Group	-1.74	1.41	-1.24	0.22
Block	0.30	0.02	18.90	<.001
Adult Group $\times$ Block	0.24	0.02	10.54	<.001
Child Group $\times$ Block	-0.33	0.02	-13.62	<.001

Table 4

Results of the mixed effects logistic regression on perceptual noise as a function of age group and block.

Fixed effects:	Estimate	Standard Error	t value	p value
(Intercept)	6.16	8.76	7.04	<.001
Adult Group	−0.26	1.27	−2.08	0.04
Child Group	0.60	1.25	4.81	<.001
Block	−4.94	2.43	−0.2	0.84
Adult Group x Block	2.37	0.00	0.07	0.95
Child Group x Block	1.08	3.48	0.31	0.76

Table 5

Results of the mixed effects logistic regression on final block accuracy as a function of age group, final block conjunctive strategy use, and perceptual noise.

Fixed effects	Estimate	SE	z value	p value
(Intercept)	1.44	0.32	4.53	<.001
Adult Group	0.42	0.42	1.00	0.32
Child Group	0.26	0.95	0.27	0.79
Non-CJ	-1.00	0.35	-2.86	<.001
Perceptual Noise	-2.98	0.91	-3.27	<.001
Adult Group x Non-CJ	0.23	0.52	0.44	0.66
Child Group x Non-CJ	-0.77	0.97	-0.79	0.43
Adult Group x Perceptual Noise	-1.46	1.27	-1.15	0.25
Child Group x Perceptual Noise	-1.54	2.74	-0.56	0.57
Non-CJ x Perceptual Noise	1.75	0.93	1.88	0.06
Adult Group x Non-CJ x Perceptual Noise	-0.04	1.48	-0.03	0.98
Child Group x Non-CJ x Perceptual Noise	2.34	2.75	0.85	0.39

Table 6

Results of the mixed effects logistic regression on auditory rule-based learning accuracy as a function of trial and Wisconsin Card Sorting Test.

Fixed effects	Estimate	Standard Error	z value	p value
(Intercept)	-2.74	7.12	-3.86	<.001
Trial	1.00	6.37	15.83	<.001
WCST	9.55	4.74	2.01	.04
Trial × WCST	2.56	4.42	5.80	<.001

Table 7

Results of the mixed effects logistic regression on conjunctive strategy use as a function of block and Wisconsin Card Sorting Test.

Fixed effects	Estimate	Standard Error	z value	p value
(Intercept)	-1.48	0.001	-857.8	<.001
Block	0.19	0.001	116.0	<.001
WCST	0.02	0.001	14.2	<.001
Block × WCST	.01	0.001	7.5	<.001