Cognition 140 (2015) 95–110

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Auditory working memory predicts individual differences in absolute pitch learning

Stephen C. Van Hedger*, Shannon L.M. Heald, Rachelle Koch, Howard C. Nusbaum

Department of Psychology, The University of Chicago, United States

ARTICLE INFO

Article history:
Received 16 June 2014
Revised 3 December 2014
Accepted 27 March 2015
Available online 20 April 2015

Keywords:
Working memory
Absolute pitch
Category learning
Perceptual learning
Expertise
Individual differences

ABSTRACT

Absolute pitch (AP) is typically defined as the ability to label an isolated tone as a musical note in the absence of a reference tone. At first glance the acquisition of AP note categories seems like a perceptual learning task, since individuals must assign a category label to a stimulus based on a single perceptual dimension (pitch) while ignoring other perceptual dimensions (e.g., loudness, octave, instrument). AP, however, is rarely discussed in terms of domain-general perceptual learning mechanisms. This is because AP is typically assumed to depend on a critical period of development, in which early exposure to pitches and musical labels is thought to be necessary for the development of AP precluding the possibility of adult acquisition of AP. Despite this view of AP, several previous studies have found evidence that absolute pitch category learning is, to an extent, trainable in a post-critical period adult population, even if the performance typically achieved by this population is below the performance of a "true" AP possessor. The current studies attempt to understand the individual differences in learning to categorize notes using absolute pitch cues by testing a specific prediction regarding cognitive capacity related to categorization – to what extent does an individual's general auditory working memory capacity (WMC) predict the success of absolute pitch category acquisition. Since WMC has been shown to predict performance on a wide variety of other perceptual and category learning tasks, we predict that individuals with higher WMC should be better at learning absolute pitch note categories than individuals with lower WMC. Across two studies, we demonstrate that auditory WMC predicts the efficacy of learning absolute pitch note categories. These results suggest that a higher general auditory WMC might underlie the formation of absolute pitch categories for post-critical period adults. Implications for understanding the mechanisms that underlie the phenomenon of AP are also discussed.

1. Introduction

Absolute pitch (AP) is often defined as the ability to name a pitch using the category of a musical note, or to produce a musical note without the aid of a reference note (e.g., Ward, 1999). The ability is reported to be remarkably rare in Western cultures, with an estimated prevalence of less than one in 10,000 individuals (e.g., Bachem, 1955; Deutsch, 2013). While the ability to name an isolated musical note might not seem to be particularly important—more akin to a party trick than a useful skill— historically AP has been viewed as a desirable ability (Takeuchi & Hulse, 1993). This is partly due to the reports that several well-known composers, such as Mozart, possessed AP (Deutsch, 2002).

Despite years of empirical research, there is still no consensus on why some individuals seem to develop AP while others do not. There is, however, considerable evidence in

---

* Corresponding author at: 5848 S. University Ave. B406, Chicago, IL 60637, United States. Tel.: +1 (810) 623 2564.

E-mail address: shedger@uchicago.edu (S.C. Van Hedger).

http://dx.doi.org/10.1016/j.cognition.2015.03.012
0010-0277/© 2015 Published by Elsevier B.V.

support of the *critical period* or *early learning hypothesis* (ELH) view of AP acquisition, which asserts that note labels must be acquired before a certain critical period of development, after which "true" [1] AP ability cannot be cultivated (for a review, see Deutsch, 2013). In support of the ELH, researchers have found that all infants are able to use absolute pitch information (Saffran & Griepentrog, 2001; Saffran, Reeck, Niebuhr, & Wilson, 2005). However, since music is generally understood in relative terms (e.g., transposing a song to a different key does not change the identity of the song), children presumably abandon absolute pitch in favor of relative pitch listening strategies, unless they explicitly learn absolute note names by a certain critical age. Also in support of the ELH, numerous studies have shown that most AP possessors report that they began musical instruction at an early age (4–6 years old) with very few AP possessors reporting that they began musical instruction at a later age (Baharloo, Johnston, Service, Gitschier, & Freimer, 1998; Chin, 2003; Deutsch, Henthorn, Marvin, & Xu, 2006; Sergeant, 1969). Finally, several training studies have found that children are able to learn absolute pitch categories faster and more accurately than adults, even if they do not reach the level of performance that is typically seen among "true" AP possessors (Crozier, 1997; Russo, Windell, & Cuddy, 2003).

The ELH seems to preclude the idea that an adult, through perceptual training, can improve their absolute pitch abilities to the extent of "true" AP possessors (see Takeuchi & Hulse, 1993). This is ostensibly because adults have missed the critical period of AP category acquisition. Put another way, the ELH would predict dichotomous absolute pitch ability in adults, with individuals either failing to remember absolute pitch information thus unable to match this information to a note category, or remembering absolute pitches in a fundamentally superior way depending on early childhood experience within a critical period window. Indeed, when using traditional note labeling tests for absolute pitch ability, performance seems to be dichotomous – delineating "true" AP possessors from non-AP possessors (e.g., Athos et al., 2007). However, there is enough variability in the population in terms of general absolute pitch ability that this assertion of a dichotomy needs to be reexamined.

The idea that there might be different levels of absolute pitch memory is not new (for a detailed description, see Bachem, 1937). In fact, a growing body of research suggests that most adults who would easily fail standard AP tests nevertheless have some long-term absolute pitch representations (albeit with higher variances than true AP possessors). Studies that research pseudo-absolute pitch (sometimes called *residual*, *implicit*, or *latent* AP) highlight the ability of non-AP possessing individuals to recognize when a familiar song is played in the correct absolute key (e.g., Schellenberg & Trehub, 2003), hum popular songs in the correct absolute key (Levitin, 1994), recognize when a highly familiar non-musical stimulus (dial tone from a landline phone) is presented at the correct or incorrect pitch (Smith & Schmuckler, 2008), and even rate isolated pitches as more pleasing if they occur less frequently in one's environment (Ben-Haim, Eitan, & Chajut, 2014). These studies suggest that most individuals have some representation of absolute pitch, even if they do not possess the ability to explicitly label tones like a "true" AP possessor. Yet, given that this type of absolute pitch knowledge is considerably more variable than "true" AP ability (as individuals typically can identify familiar songs in the correct absolute key with approximately 60–75% accuracy), it is possible that this implicit absolute pitch knowledge is fundamentally distinct from the phenomenon of "true" AP.

Even when absolute pitch ability is measured using more traditional means, such as explicitly labeling an isolated pitch with its musical note name, there is considerable variability in performance that suggests explicit absolute pitch category knowledge is not purely dichotomous. Bermudez and Zatorre (2009) found considerable evidence for an "intermediate" level of AP performance (i.e. performance that was clearly above chance, but more variable that what is commonly defined as "true" AP ability).

Finally, among "true" AP possessors (who achieve near-perfect accuracy in tests of explicit tone-label associations), there is notable variability in AP category identification with respect to both accuracy and latency (e.g., Miyazaki, 1990) that appears to reflect individual experiences listening to and labeling musical notes. For instance, AP performance appears to be more accurate for more familiar instruments (Bahr, Christensen, & Bahr, 2005; Ward & Burns, 1982), more accurate for more frequently experienced notes (Deutsch, Le, Shen, & Li, 2011), including white key notes compared to black key notes, (Takeuchi & Hulse, 1991), and even more accurate for individual notes that are used as tuning standards, such as a B-flat for a brass player (Bahr et al., 2005). Moreover, Hedger, Heald, and Nusbaum (2013) demonstrated that the tuning of AP categories in adults is malleable and dependent on environmental input, suggesting that AP categories are not crystallized and immutable after a critical period of learning, but rather can be shifted to accommodate different listening experiences. These results, taken together, suggest that regardless of the mechanisms that underlie the acquisition of "true" AP ability, learning mechanisms appear to influence the strength of particular AP note categories in a "true" AP population.

Overall, these findings across non-AP possessors, "intermediate" AP possessors, and "true" AP possessors suggest that absolute pitch recognition might be considerably more variable in the population than has been thought historically. If the acquisition and maintenance of absolute

---

[1] The use of the term "true" AP is intended to specify that AP is traditionally thought to reflect a perceptual ability that conforms to certain theoretic notions such as early acquisition. However, even high levels of AP performance show variability (Bachem, 1937), which calls into question whether there is a single "true" form of AP. Given a set of objective criteria for the classification of note labeling performance as AP, the distinction between "true" AP (canonical conformance to theoretic specification according to criteria not solely related to labeling performance) and manifest AP (labeling performance meeting all objective performance criteria) should become moot. Nevertheless, the quoted form here is intended therefore to reflect recognition that within this area of research, following training, there has been skepticism regarding improvements in note labeling performance as indicative of the same underlying mechanism.

pitch categories is conceptualized as a skill – appearing overly dichotomous because of the ways in which it is specifically tested – then it is possible that learning to associate specific pitches with musical labels (e.g., learning that a pitch of 440 Hz should be labeled as an "A") can be conceptualized as an exercise in perceptual category learning (cf. Ashby & Maddox, 2005). On its face, the explicit training of absolute pitch categories seems like a perceptual learning task (Goldstone, 1998), as individuals must learn to attend to the relevant features of a sound (i.e. pitch), while ignore features that are irrelevant for successful categorization (e.g., loudness, octave, instrument). Moreover, as Gibson and Gibson (1955) first described, individuals must engage in both differentiation (e.g., telling adjacent notes apart) and enrichment (e.g., recognizing a "C" across multiple octaves and timbres) processes.

To address whether learning absolute pitch note categories in a non-AP, adult population follows the same constraints as learning other perceptual categories, we specifically investigate one cognitive capacity that has been shown to be predictive of other cognitive processes – working memory (WM). Working memory – the higher-order cognitive ability to temporarily maintain items online (e.g., Engle, 2002) – has been shown to predict the success with which one can learn a variety of category mappings, from simple rule-based categorization (e.g., DeCaro, Thomas, & Beilock, 2008; Lewandowsky, Newell, Yang, & Kalish, 2012), to information-integration categorization, which requires individuals to integrate information across multiple perceptual dimensions (Lewandowsky et al., 2012). Presumably, the link between working memory and category learning exists because participants with higher or more efficient working memories can effectively allocate attention toward the relevant features for categorization, while allocating attention away from irrelevant features for categorization (e.g., Kane & Engle, 2000; Kruschke, Kappenman, & Hetrick, 2005; Lewandowsky, 2011).

However, from the perspective of some theories of categorization, it is not at all clear that WMC would be relevant to learning absolute pitch categories. According to the multiple memory systems (MMS) view of category learning (e.g., Ashby & Maddox, 2005), working memory does not help performance in all category learning tasks. Specifically, one prominent idea according to the MMS view of category learning is that working memory is primarily useful in learning categories with explicit rules that can be held in mind verbally (cf. DeCaro et al., 2008). Learning absolute pitch categories does not clearly fit within the specific taxonomy of rule-based category learning given that there are no clear rules to identify pitches that can be verbally explicit and useful. Since the development of absolute pitch categories does not seem to be an explicit, rule-based model of category learning, it is possible that the learning of absolute pitch categories is not affected by an individual's general auditory WMC, although it does share some characteristics with the problem of perceptual learning of phonetic categories (e.g., Nusbaum & Schwab, 1986; Schwab, Nusbaum, & Pisoni, 1985) and previous work demonstrating that perceptual learning of synthetic speech interacts with WMC (Francis & Nusbaum, 2009).

Additionally, even if the training of absolute pitch in an adult, non-AP population follows the same principles as other perceptual learning tasks, this does not mean that the same mechanisms are responsible for the development of "true" AP ability. Indeed, there are many examples of biological attributes (e.g., height) that fall along a continuum, with the tails of the distribution being represented through different mechanisms (e.g., gigantism or dwarfism). Thus, while it certainly could be the case that "true" AP appears to be a somewhat special case of category learning (especially since AP possessors often report developing categories immediately and effortlessly), there is some preliminary evidence to suggest that absolute pitch note category learning in a "true" AP population might involve similar mechanisms that subserve other forms of perceptual category learning. Deutsch and Dooley (2013) have recently demonstrated that AP possessors have a larger auditory digit span compared to non-AP possessors who were matched in age, age of musical onset, and overall music experience. This finding, while correlational, suggests that individuals may develop what is conventionally known as AP because they have a high auditory WM capacity, though the reverse is also possible (i.e. individuals first gain AP and then improve their auditory WM). If the first interpretation is correct, then it suggests that AP category acquisition and other perceptual category acquisition might be explained by similar learning mechanisms.

The present experiments were designed to investigate the degree to which post-critical period adults can explicitly learn absolute pitch categories. In particular, the question is to what extent does auditory WMC predict absolute pitch category acquisition in adults? Through measuring individual differences in the efficacy of learning absolute pitch categories, we can begin to address whether the development of absolute pitch categories in a non-AP population should be conceptualized as a difficult perceptual learning task. Specifically, if absolute pitch ability across all individuals is tied to a critical period of learning (as stated in the ELH), then we would predict that only the participants who had begun musical instruction at an early age would show any significant improvement. Indeed, previous work has found that musical experiences are specifically associated with how well a non-AP possessor can learn AP categories (Cuddy, 1968; Mull, 1925). On the other hand, if absolute pitch category acquisition can be thought of as a general perceptual category learning task, then we might expect to observe improvement across most participants, with the amount of improvement being predicted by measures of auditory WM rather than specifically musical experiences.

2. Experiment 1

Many adults (i.e. after a putative critical developmental period) have attempted to teach themselves or others AP over the past century. While the general consensus is that "true" AP cannot be taught to a previously naive adult (e.g., Deutsch, 2013; Levitin & Rogers, 2005), almost all of the efforts to learn AP have resulted in *some* improvement, with a couple of studies even claiming that individuals approached performance levels comparable to "true" AP performance after training (e.g., Brady, 1970; Rush, 1989). In the most successful studies of Brady (1970) and Rush (1989), participants' accuracy and speed at classifying isolated musical pitches was comparable to "true" AP possessors (i.e. those who had learned note names at a very young age). These scattered claims of teaching "true" AP to post-critical period adults, however, have largely been ignored or dismissed, since they either did not follow up with participants to see how the pitch categories were maintained after practice, or the participants showed some errors (such as not being able to name simultaneously presented notes) that many "true" AP possessors do not show (Takeuchi & Hulse, 1993). Yet, given the impressive variability found in "true" AP possessors with regard to instrumental timbre and pitch register (Takeuchi & Hulse, 1993), dismissing adult absolute pitch learning on the basis of particular types of errors – which might be the result of insufficient practice, rather than a fundamentally different phenomenon – might be unwarranted.

Unfortunately, there are several issues with the previous adult absolute pitch learning studies that have precluded the possibility of measuring how individual differences might interact with absolute pitch category acquisition. First, virtually all of the previous absolute pitch learning studies involved a relatively small sample size (usually just a few people, and sometimes as few as one). Second, the method of teaching absolute pitch has varied considerably, with several studies using a paradigm in which participants are played a note, identify the note, and then receive feedback (e.g., Gough, 1922; Hartman, 1954; Lundin & Allen, 1968; Vianello & Evans, 1968), and other studies using a paradigm in which participants are first taught a single pitch, and then eventually learn to discriminate this pitch from all other pitches (Brady, 1970; Cuddy, 1968). This difference in learning strategies makes trying to compare individuals across these paradigms – especially with low sample sizes – difficult.

The current study addresses these issues by standardizing the amount and nature of explicit absolute pitch category training across participants, as well as including a large enough sample size to determine which (if any) individual differences are predictive of absolute pitch category acquisition. The present study examines whether implicit note memory, as measured by a tone matching task, relates to learning absolute pitch labeling and generalization. If auditory working memory is related to learning note categories, then there should be a positive relationship between implicit note memory in pitch matching and learning note categories without a reference note.

2.1. Methods

2.1.1. Participants

Seventeen University of Chicago students participated in the experiment ($M = 20.6$, $SD = 2.6$ years old, age range: 18–26). No participants reported having absolute pitch, and all participants had a variable amount of music experience ($M = 7.4$, $SD = 4.8$ years, range: 0–14). Participants were not specifically recruited for their musical backgrounds. All participants were compensated for their participation in the experiment.

2.1.2. Materials

Participants listened to all auditory stimuli through Sennheiser HD280 studio monitor headphones. The computer screen displayed images with a 1280 × 1024 screen resolution, at a 75 Hz refresh rate. Acoustic sine waves were generated in Adobe Audition with a 44.1 kHz sampling rate and were then RMS normalized to 75 dB SPL. Instrumental notes were sampled from real instruments using the database in Reason 4.0, which is software for music production (www.propellerheads.se). The instrumental notes were also recorded in Adobe Audition with a 44.1 kHz sampling rate and were RMS normalized to 75 dB SPL. Our test for implicit note memory was run using the Psychophysics Toolbox in Matlab (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997), while our explicit pitch-labeling task was run using E-Prime (www.pstnet.com).

2.1.3. Procedure

All participants completed an implicit note memory (INM) task and an explicit pitch-labeling task. The implicit note memory task was similar to that used by Ross, Olson, and Gore (2003) to test for absolute pitch ability in nonmusicians, and has been previously used to explore auditory category memory across individuals (Heald, Van Hedger, & Nusbaum, 2014). On each trial, participants heard a brief (250 ms) sine wave target note, which was then masked by 1000 ms of white noise. Participants then had to adjust a starting note (1–7 semitones higher or lower than the target note) to try and recreate the originally heard target note. This was achieved by clicking on upward and downward arrows on the computer screen. The arrows moved the pitch either 33 or 66 cents upward or downward, depending on whether participants were clicking on the smaller arrows (33 cents) or larger arrows (66 cents). Fig. 1 shows the layout of the screen, as well as the distribution of notes. When participants believed that they had successfully recreated the original target note, they pressed a key to move onto the next trial. There were a total of four target notes (F#, G, G#, A) and eight starting notes (D, D#, E, and F below the target notes, and A#, B, C, and C# above the target notes). The entire set of stimuli spanned one octave (excluding the two microtonal steps between the highest starting note, C#, and the D from the adjacent octave), meaning there were a total of 34 notes in the series (including the microtonal traversable notes). Participants randomly heard all combinations of target note/starting notes twice, resulting in 64 trials (4 target notes × 8 starting notes × 2 repetitions). While the INM task was clearly musical in nature, as the target notes and starting notes were taken from the Western musical scale, due to the particular nature of the task we interpreted performance in terms of auditory working memory precision, since participants needed to remember the perceptual details of the target tone in the face of white noise and several intermediary tones. In

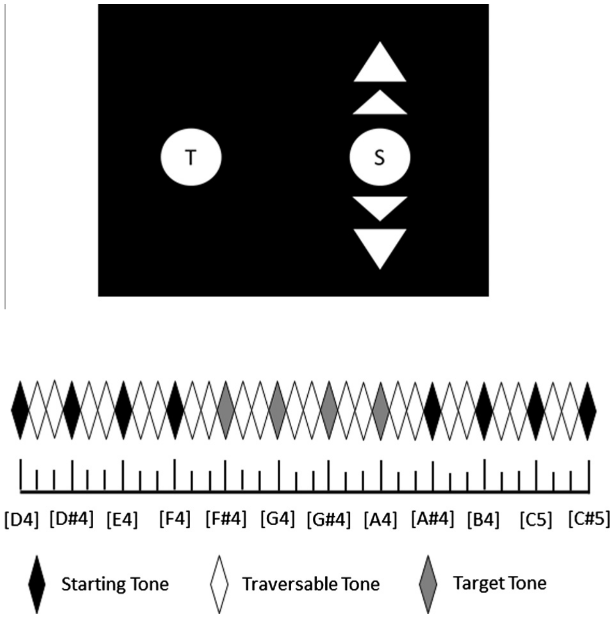

Fig. 1. Layout of the computer screen for the implicit note memory task (top), as well as a distribution of the starting tones, target tones, and traversable tones (bottom). Each step represents a 33 cent difference in pitch. Thus, the smaller, inside arrows would move the pitch by one diamond, while the larger, outside arrows would move the pitch by two diamonds.

support of our interpretation, a very similar pitch-matching task was recently used by Kumar et al. (2013) to measure the precision of pitch in auditory working memory.

The explicit pitch-labeling task consisted of three parts – pretest, training, and posttest. During pretest, participants heard an isolated piano note (1000 ms), and then attempted to identify the note by its musical note name (e.g., C or F#) by pressing a corresponding key on a computer keyboard. There were 12 possible notes spanning a one-octave range (C [4] 261.6 Hz to B [4] 493.9 Hz). Each note was presented five times for a total of 60 trials. Participants received no feedback on their answers. Moreover, for all portions of the explicit pitch-labeling task, participants heard 1000 ms of white noise and 2000 ms of 16 randomized piano tones between each trial to minimize the strategy of using relative pitch to perform the task.

For the training portion of the task, participants listened to and classified 180 piano notes (3 blocks of 60 notes per block). The procedure was identical to the pretest, except that participants received feedback on their answers. Specifically, after making their judgment, participants saw the correct label (e.g., C#) displayed on the screen, as well as re-heard the note.

The posttest was split into two parts – a rote posttest and a generalization posttest. The rote posttest was identical to the pretest, in that participants heard 60 isolated piano notes within a one-octave range, and then made their judgments without feedback. The generalization posttest consisted of 48 notes, spanning multiple octaves and instruments. Specifically, in the generalization posttest, participants heard 12 piano notes from the original octave distribution (C [4] 261.6 Hz to B [4] 493.9 Hz), 12 piano notes from the adjacent, higher octave (C [5] 523.3 Hz to B [5] 987.8 Hz), 12 acoustic guitar notes from the original octave distribution (C [4] 261.6 Hz to B [4] 493.9 Hz), and 12 acoustic guitar notes from the adjacent, lower octave (C [3] 130.8 Hz to B [3] 246.9 Hz). The reason for including the generalization posttest is that we wanted to measure whether the note category learning that occurred over the course of the experiment generalized to frequencies and timbres that were not specifically trained. Furthermore, during training on a limited range of musical notes, it is possible that participants learned the boundaries of the distribution (e.g., that C was the lowest note), and thus relied on non-absolute cues to make their judgments. By introducing multiple instruments and octaves, we made it considerably more difficult to rely on non-absolute cues to successfully perform the task.

After the implicit note memory task and the explicit pitch-labeling task, participants filled out a music experience questionnaire. Participants were then debriefed and compensated with either money or course credit.

2.2. Results

2.2.1. Rote and generalized learning

We first assessed whether participants showed any improvement in note classification as a function of training in the explicit AP category learning task through constructing a repeated measures analysis of variance, with test type (pretest, rote posttest, generalization posttest) as repeated factors. The overall analysis of variance was significant [$F(2,32) = 22.17$, $p < 0.001$]. Using a Fisher LSD post hoc test to compare mean differences among the three tests, we found that individuals were significantly more accurate at identifying notes after training, correctly identifying 13.7% (SD = 10.8%) of one-octave piano notes in the pretest, and correctly identifying 36.2% ($SD = 19.4\%$) of one-octave piano notes after training in the rote posttest [$t(16) = 5.35$, $p < 0.001$]. For the generalization posttest, which consisted of octaves and timbres that were not specifically trained, performance decreased relative to the rote posttest [$t(16) = 6.26$, $p < 0.001$] with participants only identifying 21.7% ($SD = 15.3\%$) of notes correctly. However, performance in the generalization posttest was still significantly above pretest performance [$t(16) = 2.31$, $p < 0.05$], which is notable considering the pretest was ostensibly easier, as it only tested one octave of piano notes. Moreover, as assessed through a one-sample *t*-test, both the rote (36.2%) [$t(16) = 5.91$, $p < 0.001$] and generalization (21.7%) [$t(16) = 3.60$, $p = 0.002$] posttests were significantly above chance performance (represented as 1/12, or 8.33%). Taken together, these results strongly demonstrate that participants showed significant improvements in both rote and generalized learning as a function of training.

2.2.2. Implicit note memory performance

To examine participants' performance on the INM task, we took the absolute value of the difference between a participant's final location to which they moved the probe tone in pitch space and the true target note. For example, if a participant's target note was [A4] and their final location in recreating this [A4] was [A#4], they would be three 33-cent steps from the true location, and thus receive a score of "3" on that particular trial. We collapsed across all trials, calculating a single INM score per participant.

Overall, participants were relatively good at adjusting a starting probe tone to match a target tone, as they were on average only 1.31 steps (approximately 40 cents) away from the true target note. This difference was still significantly above zero, suggesting that participants reliably demonstrated error in recreating the target tone [$t(16) = 8.68$, $p < 0.001$]. Furthermore, there was considerable individual variability in performance, from an average of 0.45 steps (approximately 15 cents) away from the target note, to an average of 2.36 steps (approximately 78 cents) away from the target note. Given that the just noticeable difference (JND) for sine waves within the tested frequency range is approximately 10 cents (e.g., Kollmeier, Brand, & Meyer, 2008), the highest performing individuals' average deviation was higher than the difference limen in auditory pitch. The individual differences in INM performance were significantly correlated with overall musical instruction (operationalized as the number of years spent playing one's primary instrument) [$r = -0.42$, $n = 17$, $p < 0.05$], as well as the age of music onset [$r = 0.64$, $n = 17$, $p < 0.01$], which is consistent with prior work showing that music experience is associated with an enhancement in domain-general auditory processes (e.g., Kraus & Chandrasekaran, 2010).

2.2.3. Predicting AP learning with working memory

In order to test our hypothesis that auditory working memory or the age of music onset would predict how well an individual acquired explicit absolute pitch categories, we constructed a generalized mixed-effects model (e.g., Baayen, Davidson, & Bates, 2008) with a binomial link. Specifically, we treated INM score and the age of music onset as fixed effects, while we treated participant and note (stimulus) as random effects. Explicit absolute pitch category learning (measured as a proportion of correct answers) was our dependent variable, while INM score and age of music onset were our predictor variables. We operationalized explicit absolute pitch category learning by looking at posttest performance on the block of multiple instruments and timbres (generalization posttest). The reason we specifically used performance on the generalization posttest (rather than the rote posttest) is because it provided a more stringent test of absolute pitch category knowledge, as participants needed to generalize beyond the specific stimuli upon which they were trained in order to succeed. Moreover, we included age of music onset as a predictor variable since previous research suggests that individuals who have specifically early experience with note labels might show the most improvement in explicit absolute pitch note category learning (cf. Crozier, 1997). We did not include amount of musical instruction in our model, since age of music onset and overall musical instruction were highly correlated [$r = -0.73$, $n = 17$, $p < 0.001$] and thus might introduce issues of multicollinearity. Another theoretical reason for not including both age of music onset and overall musical instruction in the same model is because these two measures are presumably tapping into the same broad construct of musicianship. Thus, even though the INM task was significantly correlated with musical experience (operationalized as age of music onset and overall musical instruction), we included it in our model because it was meant to assess implicit auditory working memory, which is related to – but dissociable from – musicianship. Indeed, previous research using similar tests for implicit note memory have clearly demonstrated that performance can be interpreted in terms of precision of auditory working memory (Kumar et al., 2013), and performance is not necessarily tied to specifically musical experiences (e.g., Ross et al., 2003).

We first constructed simple models to assess whether INM score or the age of music onset would predict explicit absolute pitch category learning in isolation. Indeed, we found that both INM score [$\beta = -1.065$, $SE = 0.254$, $p < 0.0001$] and the age of music onset [$\beta = -0.115$, $SE = 0.044$, $p < 0.01$] significantly predicted explicit AP category learning in isolation. Moreover, the age of music onset significantly predicted performance on the INM task [$\beta = -0.361$, $SE = 0.111$, $p < 0.01$]. In a combined model, however, INM score was the only significant predictor of explicit AP category learning. The age of music onset failed to significantly predict explicit AP category learning in the model including INM score (see Table 1). The adjusted *R*-squared value for the model including age of music onset and INM score was 0.388, meaning that we were able to account for 38.8% of the variance in absolute pitch learning using just two variables.

This relationship between INM score, the age of music onset, and explicit absolute pitch category learning suggests that perhaps the relationship between the age of music onset and explicit AP category learning was being mediated by auditory working memory. Indeed, a Sobel test for mediation revealed that our auditory working

Table 1
Multiple regression output from a generalized mixed-effects model, with age of music onset and INM score as fixed effects, and participant and musical note (stimulus) as random effects. While age of music onset significantly predicted explicit absolute pitch category learning in isolation, it fails to do so in a model including INM score.

Fixed effects	Estimate	Standard error	*Z*-value	*Pr*(>|*z*|)
(Intercept)	−0.029	0.403	−0.072	0.943
INM score	−0.935	0.325	−2.875	0.004
Age of music onset	−0.030	0.048	−0.627	0.530

memory measure – INM – was significantly mediating the relationship between age of music onset and absolute pitch category learning [$t = -2.16$, $SE = 0.16$, $p = 0.03$]. This mediation relationship is represented in Fig. 2.

Given our relatively low sample size ($n = 17$), we also assessed mediation through bootstrapping procedures. The index of mediation (see Preacher & Hayes, 2008) was calculated for 10,000 bootstrapped samples. The bootstrapped index of mediation was −0.14 and the 95% confidence interval did not include zero (−0.20, −0.08). Thus, using both Sobel's Test for mediation as well as bootstrapping techniques, we found evidence that auditory working memory was significantly mediating the relationship between age of music onset and explicit absolute pitch category learning.

2.2.4. Retention of absolute pitch categories

Absolute pitch training studies are generally criticized for not retesting participants after training ceases (see Takeuchi & Hulse, 1993) to determine the rate at which category memory is lost. Since the performance of "true" AP possessors does not seem to significantly change over a short-term time course, critics of adult absolute pitch category learning studies claim that while non-AP possessors might appear behaviorally indistinguishable from non-AP possessors after sufficient training, active training is required to maintain AP-like performance.

There are, however, a number of concerns with this reasoning. First, performance on note category tests for "true" AP possessors has been shown to vary based on numerous factors, including the age of the participant (Athos et al., 2007), the menstrual cycle of (female) participants (Wynn, 1973), and the intonation of the immediately preceding musical context within a single laboratory session (Hedger et al., 2013). Thus, the claim that "true" AP possessors do not significantly vary in their note judgments on a short time course does not appear to be entirely accurate. Second, just because previous absolute pitch training studies have not retested their participants after training does not mean that the participants in the studies have lost all of their learned note category information. Indeed, Brady (1970) reported that he was able to accurately identify musical notes within one semitone five months after his training ended, though he only tested himself on five notes. Moreover, given the generally low sample sizes for AP training studies, drawing definitive conclusions from retesting one or two participants is difficult.

We were able to retest 6 of our 20 participants, from five to seven months after the previously described training session [$M = 184$ days, $SD = 22$ days]. With respect to INM performance, we obtained a representative sample from the experiment [$M = 1.28$, $SD = 0.68$ for retested participants, $M = 1.33$, $SD = 0.66$ for non-retested participants, $t\,(15) = 0.16$, $p > 0.8$]. The six retested participants, however, had marginally more musical experience compared to the non-retested participants [$M = 9.50$, $SD = 3.89$ years for retested participants, $M = 5.32$, $SD = 4.97$ years for non-retested participants, $t\,(15) = -1.78$, $p = 0.10$]. Age of music onset was not significantly different between the retested and non-retested groups [$p = 0.12$]. No retested participant reported actively rehearsing or retraining note categories since the original learning session. During the retest, participants completed an abridged version of the rote posttest, in which they heard every note four times (48 trials), as well as the full version of the generalization posttest (48 trials). The delayed generalization posttest was the exact same as the generalization posttest immediately following training (i.e. consisting of the same pitch ranges and timbres). Participants did not receive feedback on their performance.

The results are displayed in Fig. 3. To assess performance loss, we constructed a 2 × 2 repeated measures analysis of variance, with test type (rote posttest, generalization posttest) and time point (immediate, delay) as repeated factors. We found a main effect of time point [$F\,(1,5) = 7.93$, $p = 0.04$], with participants losing 11.1% (*SD:* 9.6%) from the immediate posttests to the delayed posttests. Additionally, we found a main effect of test type [$F\,(1,5) = 7.63$, $p = 0.04$], with participants performing on average 15.7% (*SD*: 13.9%) worse on the generalization

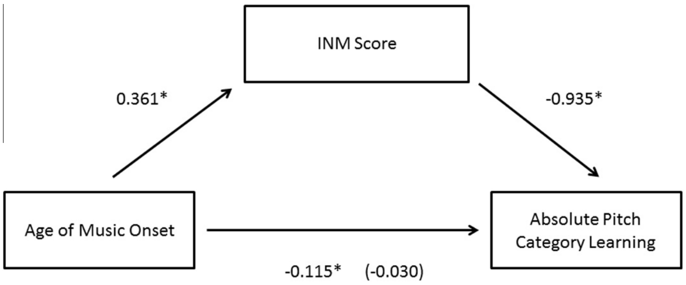

Fig. 2. Mediation relationship between age of music onset, INM score, and absolute pitch category learning (Experiment 1). While the age of music onset significantly predicts explicit absolute pitch category learning in isolation, ostensibly providing support for a critical period model of AP learning, auditory working memory (measured through INM score) significantly mediates this relationship. $^{*}p < 0.05$.

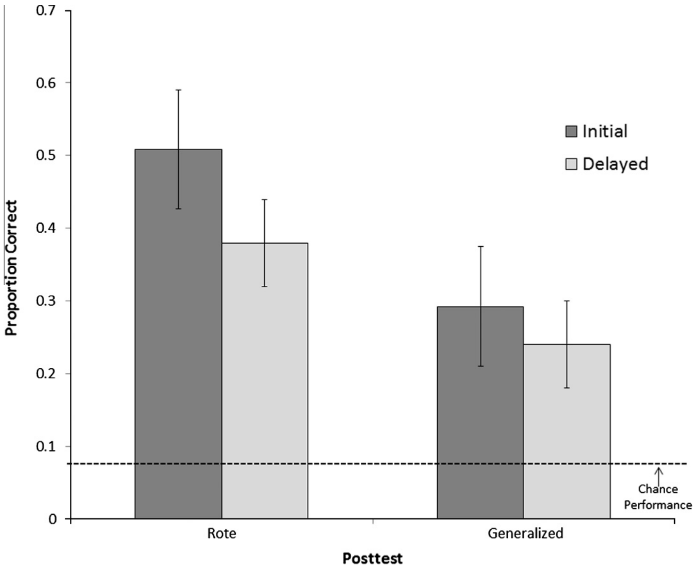

Fig. 3. Initial and delayed posttest scores for 6 of the 17 re-tested participants in Experiment 1. The dotted line represents chance performance. Error bars represent ±1 SEM.

posttests than on the rote posttests. Despite the significant overall performance loss between the immediate posttests and the delayed posttests, participants were still above chance (1/12 or 8.3%) in the delayed rote posttest, as determined by a one-sample *t*-test [$t\ (5) = 5.80$, $p < 0.01$]. For generalized category learning, performance decreased approximately 5 percentage points, from 29.2% (*SD*: 21.5%) in the immediate generalization posttest to 24.0% (*SD*: 18.0%) in the delayed generalization posttest. This loss, however, was not statistically significant [$t\ (5) = 1.86$, $p = 0.12$]. Moreover, performance in the delayed generalization posttest was marginally above chance [$t\ (5) = 2.13$, $p = 0.08$]. Thus, after an average span of 184 days without training, we found evidence that participants lost approximately 11 percentage points of learning, though performance was still significantly above chance in the rote posttest and marginally above chance in the generalization posttest. Even though the results from the generalization posttest are not as strong as the results from the rote posttest, they are particularly compelling as the generalization posttest contained several stimuli that never were trained (i.e. participants never received feedback). Thus, while this retest must be interpreted with caution given the low sample size, the notion that any performance gains made by adults attempting to learn absolute pitch categories will be completely lost without explicit rehearsal may not be completely accurate.

2.3. Discussion

Experiment 1 was designed to assess whether pitch matching performance (as a test of auditory working memory) and a measure of music experience might contribute to learning absolute pitch labels in a rapid (single session) perceptual learning study. Based on the theory that absolute pitch ability lies on a continuum and can be tuned through perceptual learning, we hypothesized that auditory working memory ability would predict explicit note category learning, since working memory capacity has been implicated in explaining performance variability in a wide variety of perceptual tasks, (e.g., Daneman & Carpenter, 1980), as well as in explicit categorization tasks (Lewandowsky, 2011). If, however, absolute pitch category learning to any extent is relegated to a critical period in development, we would not predict that auditory working memory ability should be related to absolute pitch category learning, as learning might not be possible in an adult population.

Our results provide evidence for a more general perceptual learning theory of absolute pitch category acquisition in an adult population. While this more general mechanism might not underlie the acquisition of “true” AP ability, it is notable that we observed a significant improvement in absolute pitch categorization for non-AP possessors in a single learning session – both for rote and generalized learning. Moreover, we found that absolute pitch category learning in an adult non-AP population was predicted by auditory working memory ability, as measured by an implicit note memory task similar to that used by Ross et al. (2003). While the age of music onset significantly explained absolute pitch category learning in isolation, it failed to retain significance in the model that included our working memory measure. This suggests that the age at which individuals first learned musical note names was less important than their more generalized ability to integrate many perceptual features of a sound into a single category representation.

Our finding that auditory working memory mediates the relationship between age of music onset and absolute pitch learning in non-AP adults fits with previous research that demonstrates that early musical experiences shapes more general auditory processing abilities, which in turn leads to better auditory working memory (for a review of the generalization of music training to other domains, see Kraus & Chandrasekaran, 2010). In this sense, early musical exposure would not be necessary for absolute pitch category learning per se, but it might facilitate the acquisition of absolute pitch categories through strengthening more general auditory processes. On the other hand, it is also possible that individuals who were born with higher auditory processing abilities were more drawn to music, and consequently began musical instruction at an early age. Unfortunately, the correlational nature of the relationship between age of music onset and INM score in the present study precludes making causal inferences.

The category learning we observed in the current experiment also appears to be relatively robust, as a subsection of participants who were brought in between five and seven months after the initial learning session were still significantly above chance performance in the rote posttest and marginally above chance performance in the generalization posttest. This degree of stability in category representations (without active training) challenges the prevailing view on AP acquisition – that is, without constant rehearsal and active training, any gains made in absolute note identification by a non-AP possessor will be lost (Takeuchi & Hulse, 1993). Interestingly, these findings are remarkably similar to learning patterns seen in synthetic speech training – that is, participants generally improve about 45% from pretest to posttest, and do not appear to lose information even after considerable delays (6 months) in generalization retesting (Schwab et al., 1985). Given our reduced sample size for retesting participants' note categories, however, our results require further empirical substantiation.

3. Experiment 2

While Experiment 1 provides empirical evidence that more general auditory – rather than specifically musical – abilities predict absolute pitch category learning in a rapid perceptual learning study, there are still unanswered questions that warrant further investigation. For example, the implicit note memory task used in Experiment 1 was largely musical in foundation, since the target notes were derived from the Western musical scale. Moreover, given that the INM task required participants to maintain fine pitch differences in working memory, it is possible that the individual differences in INM performance were inherently tied to music experience, as difference limens for pitch have been shown to vary as a function of musical experience (e.g., Kishon-Rabin, Amir, Vexler, & Zaltz, 2001). Thus, given the intricate link between the INM task and music experience, it is unclear whether the individual differences observed in the implicit note memory task were the result of music specific processing mechanisms, or whether performance on the implicit note memory task correlates well with other (non-musical) measures of auditory working memory. If music experience shapes *general* aspects of auditory working memory, which in turn helps individuals learn absolute pitch categories, then a non-musical test of auditory working memory should also mediate the relationship between music experience and absolute pitch category learning. This issue is addressed in the second experiment by using an auditory n-back task with speech stimuli as our test of working memory. This specifically addresses whether there is a general auditory WMC that relates to the ability to learn to recognize musical notes without a reference note.

3.1. Methods

3.1.1. Participants

Thirty University of Chicago students, staff, and community members participated in the experiment ($M = 22.0$, $SD = 4.2$ years old, age range: 18–32, 19 male). One participant reported that they were "unsure" whether or not they had absolute pitch. When analyzing this participants' data, it became clear that they possessed some form of absolute pitch, as they performed with 97% accuracy post-training and consistently misclassified notes by 2 semitones in the pretest (70.1% were misclassified by exactly 2 semitones, 23% were misclassified by exactly one semitone, and 6% were classified correctly). Even in the pretest, no note was misclassified by more than 2 semitones. We thus omitted this participant from all analyses, leaving twenty-nine participants in our analysis. All remaining participants had a variable amount of music experience ($M = 4.6$, $SD = 6.0$ years, range: 0–26). Participants were not specifically recruited for their musical backgrounds. All participants were compensated for their participation in the experiment.

3.1.2. Materials

Participants listened to all auditory stimuli through Sennheiser HD280 studio monitor headphones. The computer screen displayed images and text with a 1280 × 1024 screen resolution, at a 75 Hz refresh rate. Instrumental notes were sampled from real instruments using the database in Reason 4.0, which is software for music production (www.propellerheads.se). The instrumental notes were also recorded in Adobe Audition with a 44.1 kHz sampling rate and were RMS normalized to 75 dB SPL. All parts of the experiment were coded and run using E-Prime (www.pstnet.com).

3.1.3. Procedure

The procedure was nearly identical to the procedure in Experiment 1, with the exception that participants completed auditory n-back (ANB) task – rather than the INM task – prior to participating in the explicit absolute pitch category learning task.

The auditory n-back task required participants to actively monitor a string of spoken letters, pressing a button labeled "Target" if the currently spoken letter matched the letter presented *n* trials previously, and pressing a button labeled "Not Target" if the currently spoken letter did not match the letter presented *n* trials previously. All participants completed an auditory 2-back and an auditory 3-back task (in that order). Both the auditory 2-back and 3-back consisted of 90 total trials (three runs of 30 spoken letters). Letters were spoken one-at-a-time, with an interstimulus-interval of 3000 ms. Targets occurred one-third of the time, while non-targets occurred two-thirds of the time. Before the 2-back and 3-back, participants completed a practice round of 30 trials to familiarize themselves with the task.

The explicit AP category learning task was virtually identical to the one used in Experiment 1. During the pretest portion of the task, participants heard 1000 ms piano tones, ranging from middle C [C4] to the B above middle C [B4], presented in a randomized order. Each of the 12 notes was presented 4 times each, resulting in 48 total trials. Participants were then trained on these same 12 piano notes for 120 trials (12 notes × 5 repetitions × 2 blocks), during which they received both auditory and visual feedback on their responses. Then, participants underwent a test of rote learning, during which they classified the same 12 piano notes five times each in a randomized order (receiving no feedback). Finally, participants underwent a test of generalized learning, during which they classified 48 notes that spanned beyond the particular timbre and octave range that was trained (for details, see the Procedure section of Experiment 1). All classified notes (during pretest, training, rote posttest, and generalization posttest) were separated by 1000 ms of white noise and 2000 ms of scrambled piano notes to minimize the ability to use relative pitch on the task.

After the ANB task and the explicit absolute pitch learning task, participants filled out a music experience questionnaire. Participants were then debriefed and compensated with either money or course credit.

3.2. Results

3.2.1. Rote and generalized learning

Similar to Experiment 1, we first assessed whether participants showed any improvement in note classification as a function of training in the explicit labeling portion of the experiment. To assess this, we constructed a repeated measures analysis of variance with test type (pretest, rote posttest, generalization posttest) as repeated factors. The overall analysis of variance was significant [$F(2,56) = 14.01$, $p < 0.001$], suggesting that at least one of the tests was significantly different from one or more of the other tests. Using a Fisher's LSD post hoc test, we found that participants significantly improved from the pretest – in which they correctly identified 10.9% (*SD:* 14.5%) of notes – to the rote posttest, in which they correctly identified 25.8% (*SD*: 22.3%) of notes [$t(28) = -4.11$, $p < 0.001$]. Performance in the generalization posttest was significantly worse than performance in the rote posttest, with participants correctly identifying 15.4% (*SD*: 12.8%) of notes [$t(28) = 4.73$, $p < 0.001$]. Despite this significant difference between performance on the rote posttest and performance on the generalization posttest, participants performed marginally better in the generalization posttest compared to their pretest performance [$t(28) = 1.80$, $p = 0.09$], which is notable considering the pretest was ostensibly easier (as it only contained notes from a single octave and a single timbre). Moreover, a one-sample *t*-test showed that performance on both the rote posttest [$t(28) = 4.21$, $p < 0.001$] and the generalization posttest [$t(28) = 2.98$, $p = 0.006$] were significantly above chance performance (1/12, or 8.3%). These results clearly demonstrate that participants showed significant improvements in both rote and generalized learning as a function of training.

3.2.2. Auditory n-back performance

We calculated auditory n-back performance using signal detection theory (e.g., Macmillan & Creelman, 1991). Specifically, we calculated the proportion of "hit" trials (correctly responding that the currently spoken letter was presented n-letters previously) and the proportion of "false alarm" trials (incorrectly responding that the currently spoken letter was presented n-letters previously). If a participant received a proportion of 1 or 0 (e.g., by scoring 30 out of 30 hits or 0 out of 60 false alarms), we calculated a proportion using the formula $((n * 2) \pm 1)/(t * 2)$, where n equals the total number of hits or false alarms, and t equals the total number of trials. For example, a subject who scored a perfect 30 out of 30 hits would receive the proportion $((30 * 2) - 1)/(30 * 2)$, or 59/60. This was done to obtain an actual *z*-score (as probabilities of 1 and 0 would correspond to *z*-scores of ∞ and $-\infty$, respectively). We then *z*-transformed and subtracted the false alarm proportion from the hit proportion to obtain a *d*-prime score for each participant. Using the correction procedure for probabilities of 1 and 0, a perfect subject (30 out of 30 hits and 0 out of 60 false alarms) would obtain a *d*-prime score of 4.52. Participants' *d*-prime scores clearly reflected their ability to detect targets for both the 2-back task [$d' = 3.35$, $SD = 0.86$, $p < 0.01$] and the 3-back task [$d' = 2.27$, $SD = 1.01$, $p < 0.01$].

3.2.3. Predicting absolute pitch learning with working memory

To assess whether we could explain the observed variance in absolute pitch category learning, we first constructed simple, mixed effects regression models using just n-back score or the age of music onset to predict AP category learning. Similar to Experiment 1, absolute pitch category learning was operationalized by looking at performance on the generalization block of the posttest. Participant and note (stimulus) were treated as random effects. We then included both predictor variables in the same model, to look at whether there was any evidence that working memory was once again mediating the relationship between age of music onset and absolute pitch learning (as was the case in the first experiment).

In a regression model, age of music onset was marginally predictive of absolute pitch category learning [$\beta = -0.037$, $SE = 0.021$, $p = 0.08$]. Moreover, age of music onset significantly predicted performance in the n-back task [$\beta = -0.059$, $SE = 0.020$, $p < 0.01$], which is consistent with the idea that musical training can enhance auditory working memory (Parbery-Clark, Skoe, Lam, & Kraus, 2009). Auditory n-back also predicted absolute pitch category learning in isolation [$\beta = 0.474$, $SE = 0.183$, $p < 0.01$]. However, in a multiple regression model that included both auditory n-back and age of music onset as predictor variables, age of music onset no longer predicted absolute pitch category learning [$\beta = -0.013$, $SE = 0.024$, $p > 0.5$], while auditory n-back performance – even when controlling for the age of music onset – significantly predicted absolute pitch category learning [$\beta = 0.413$, $SE = 0.211$,

$p = 0.05$]. The adjusted R-squared value of the model including both age of music onset and auditory n-back score was 0.237, meaning that we were able to account for 23.7% of the variance in absolute pitch learning using just two variables.

The fact that the age of music onset significantly explained variance in absolute pitch category learning in isolation, but failed to do so in a model that included auditory n-back score suggests that auditory n-back was perhaps mediating the relationship between age of music onset and absolute pitch category learning. A Sobel test for mediation was marginally significant [$t = -1.63$, $SE = 0.014$, $p = 0.10$]. This mediation relationship is represented in Fig. 4. Given our relatively low sample size, we also used a bootstrapping method for assessing mediation. With 10,000 bootstrapped samples, the index of mediation (Preacher & Hayes, 2008) was −0.052, with the 95% confidence interval not including zero (−0.074, −0.030). Thus, across two methods of assessing mediation, we found converging evidence that auditory working memory was once again mediating the relationship between the age of music onset and absolute pitch category learning.

3.3. Discussion

The purpose of Experiment 2 was to extend our findings from Experiment 1 using a different measure of auditory working memory (the auditory n-back). Our results in the current experiment are largely consistent with the first experiment – that is, while both age of music onset and auditory working memory predict individual variability in absolute pitch category learning in isolation, the relationship between age of music onset and absolute pitch category learning is mediated by auditory working memory. These results can be interpreted in a "domain-general enhancement" framework, in which early, extensive musical instruction potentially shapes general aspects of auditory working memory and selective attention, which in turn are important constructs in training absolute pitch ability in an adult population. In this sense, early musical training matters insofar as it helps explain general differences in auditory processing.

The selection of the auditory n-back task as our working memory measure requires some justification, since recent research has begun to question the construct validity of the n-back task (Kane, Conway, Miura, & Colflesh, 2007). Specifically, recent work is suggesting that the n-back task does not correlate very strongly with other working memory measures, and n-back might actually be more related to fluid intelligence (see Conway, Kane, & Engle, 2003) than working memory per se. We chose to use the auditory n-back because preliminary investigations in our lab suggest that performance on the INM task is significantly correlated with both the auditory 2-back task and the auditory 3-back task [$r = -0.51$ for both 2-back and 3-back, $n = 20$, $p = 0.01$]. We did not specifically test how our INM task correlated with other working memory measures (e.g., RSPAN, OSPAN, reverse digit span), but we would predict that working memory tests that share variance with the INM task would likely also predict individual differences in absolute pitch category learning. Nevertheless, it is notable that the auditory n-back– which does not directly require fine pitch memory and is a non-musical task – significantly predicts absolute pitch category learning.

Finally, while the mediation analysis from the current experiment largely converged with the results from Experiment 1, some caution must be exercised in interpreting this relationship. Importantly, the age of music onset only marginally predicted absolute pitch category learning in the current experiment ($p = 0.08$), which would technically stop any further testing of mediation as the p-value was greater than our alpha cutoff (0.05). However, the fact that age of music onset was marginally significant, combined with the results from Experiment 1, in which age of music onset significantly predicted absolute pitch category learning in isolation ($p < 0.01$), we continued with the tests for mediation. Moreover, Kenny and Judd (2014) have suggested that the causal variable (age of music onset) need not be significantly correlated with the outcome variable (generalized absolute pitch learning), as the lack of a significant correlation between these two variables often reflects low power. Thus, while the evidence for auditory working memory mediating the relationship between age of music onset and absolute pitch category learning is perhaps less direct compared to Experiment 1, we believe there is still converging evidence across both experiments for the role of domain-general auditory processes in learning absolute pitch categories.

There are several possible reasons why the mediating relationship of working memory in the current experiment may have been weaker than the observed relationship in

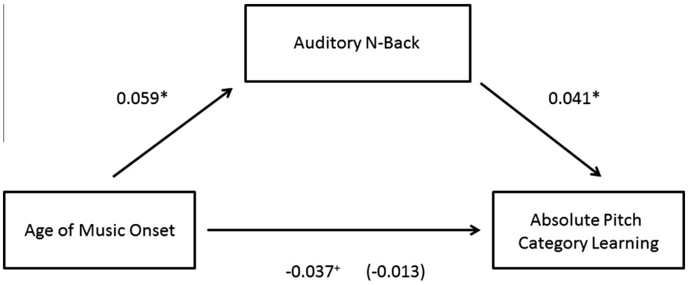

Fig. 4. Mediation relationship between age of music onset, auditory n-back performance, and absolute pitch category learning (Experiment 2). While the age of music onset significantly predicts explicit absolute pitch category learning in isolation, auditory working memory (measured through n-back) mediates this relationship. $^{*}p < 0.05$, $^{+}0.05 \leqslant p \leqslant 0.10$.

Experiment 1. First, there were significant differences between the participants in Experiments 1 and 2 with regards to music experience. Specifically, participants from Experiment 1 reported starting musical instruction at a significantly earlier age (9.6 years old) compared to participants from Experiment 2 (14.1 years old, $p = 0.02$). Additionally, 15 of the 17 participants in Experiment 1 reported at least some active instruction on a musical instrument, while only 18 of the 29 participants in Experiment 2 reported at least some active instruction on a musical instrument, which was a significant difference using Barnard's Exact Test ($p = 0.04$). These differences do not inherently explain the weaker mediating relationship in Experiment 2, though the compressed variability in musical instruction potentially explains why the overall absolute pitch learning in the current experiment was slightly lower than the learning in Experiment 1. Second, the musical nature of the INM used in Experiment 1 could have contributed to its stronger mediating role compared to the auditory n-back, especially since difference limens in pitch discrimination vary as a function of music experience (e.g., Kishon-Rabin et al., 2001). Indeed, the relationship between the age of music onset and the INM task ($r = 0.63$) was stronger than the relationship between the age of music onset and the auditory n-back task ($r = -0.50$). This increased shared variance between music experience and working memory in Experiment 1 thus could have contributed to the stronger mediating relationship between age of music onset, working memory, and absolute pitch category learning.

4. General discussion

Training absolute pitch categories in an adult population has been met with a great deal of skepticism over the past several decades. While most empirical studies report improvements in absolute pitch category identification as a function of training, very few studies have claimed to teach "true" AP with unqualified success to any post-critical period adult. With a lack of consistent evidence that "true" AP can be trained in an adult population, the question becomes whether null results should be interpreted as a true inability for adults to acquire absolute pitch categories, or whether null results reflect insufficient training (either in terms of quality or duration).

The current set of studies cannot directly address whether post-critical period adults can gain absolute pitch ability that is comparable to "true" AP ability (as our training paradigm was only a single laboratory session, and no participant reached a level of performance that is typically seen in a "true" AP population). Consequently, the present results cannot directly comment on the underlying mechanisms of the phenomenon of "true" AP. However, our finding across two studies that auditory working memory can explain the success of non-AP possessors learning absolute pitch categories supports the notion that intermediate levels of absolute pitch ability (operationalized as significantly above chance, but significantly below the level of performance typically observed among "true" AP possessors) might be best conceptualized as a domain-general perceptual learning task, rather than a specifically musical ability. To bridge the gap between the current set of studies and the relationship between auditory WM in a "true" AP population, there are two types of studies that would be important to run. First, extrapolating our results would suggest that post-critical period adults with unusually large auditory working memories –given the right amount and nature of training – can explicitly learn note-label mappings with the same success as a "true" AP possessor. While this type of study might be difficult to execute, conceptually speaking a positive result of this nature would suggest that the phenomenon of "true" AP perhaps does not operate under a critical period of development. Second, even though "true" AP ability is typically characterized as near-perfect accuracy at labeling isolated musical notes, this does not mean that there is no variability among a "true" AP population, especially for less frequently experienced timbres, such as sine tones (e.g., see Athos et al., 2007). Future studies thus might address whether individual differences within a "true" AP population can also be explained using a general measure of auditory working memory. If so, this would suggest that the fidelity of absolute pitch representations across all individuals might be intricately linked to general aspects of auditory working memory capacity.

The fact that both the INM and the auditory n-back task mediated the relationship between music experience and absolute pitch category learning is notable, especially since we found that the INM only shared around 25% of shared variance between the INM and auditory n-back task. Given the differences between these two tasks, the question becomes: why do both the INM task and the auditory n-back task significantly predict variance in absolute pitch category learning? One possible explanation is that while the INM task and the auditory n-back task reflect auditory working memory, they might represent different *aspects* of auditory working memory. Recent working memory research (e.g., Awh, Barton, & Vogel, 2007; Xu & Chun, 2006) has begun to distinguish components of working memory, such as *quality* (perceptual resolution) versus *quantity* (the number of objects able to be maintained). This distinction of quality and quantity in working memory fits nicely with the differences between the INM task and the auditory n-back task. For instance, the INM task requires the maintenance of just a single perceptual object, though successful performance necessitates that participants maintain a high perceptual resolution of this object. In this sense, *quantity* of storage capacity in WM is not necessary, but maintaining a high *quality* object in auditory WM is critical for successful performance. On the other hand, the auditory n-back requires the constant maintenance and updating of an auditory buffer, though given that the auditory objects can be easily categorized (as they are letters), and given that successful performance does not require maintaining any of the perceptual details of the spoken letters, it is likely that auditory n-back was tapping into the *quantity* aspect of WM. Both aspects of WM can be thought of as important in a perceptual category learning task such as absolute pitch learning, as one must not only maintain a high resolution representation in working memory to learn the necessary acoustic cues for categorization, but they must also hold several objects in working memory at once, in an effort to differentiate potentially similar categories (e.g., differentiating a C from a C#).

The phenomenon of "true" AP is sometimes characterized as a two-step process. First, individuals must have enhanced tonal memory, or fixed pitch-chroma categories, and second, individuals must have cultural labels (e.g., F#) that can be easily retrieved to assign to these fixed categories (e.g., Levitin, 1994; Ross, Gore, & Marks, 2005; Zatorre, 2003). Our current studies shed light on this two-step model of understanding AP. Based on our results, we would suggest that the first "step" in AP ability possibly reflects an individual's auditory working memory ability – the better the auditory working memory, the better an individual is able to form a representation based on one attribute (pitch) out of many possible attributes (timbre, loudness etc.). This notion situates higher working memory individuals to more effectively learn explicit note-label mappings (the second step of AP), as we observed across two experiments.

The idea that auditory working memory might play an important role in learning absolute pitch offers a possible causal interpretation of Deutsch and Dooley (2013), who found an enhanced auditory digit span in AP possessors compared to musically matched controls. It is possible that individuals who have unusually large auditory working memory capacity develop what is known as AP, rather than the other way around (i.e. AP consequently leads to a higher auditory working memory). While an early critical or sensitive period might seem necessary for developing "true" AP ability, it is also possible that early musical experiences further shape general auditory processes (e.g., Kraus & Chandrasekaran, 2010), including auditory working memory, which ultimately predict the efficacy of learning and using absolute pitch information. Indeed, this interpretation would also be consistent with the intriguing reports of adults who have little to no musical experience, but seem to use absolute pitch in a comparable way as "true" AP possessors in tasks that do not assume musical knowledge (Ross et al., 2003). In the model we are proposing, these rare, non-musical adults who seem to have AP, likely have extremely high executive functioning in the auditory domain, which allows them to selectively attend to pitch information and hold it in working memory even when presented with several interfering tones. In other words, these individuals are remarkably adept at the first step of "true" AP ability.

While the present studies shed light on individual differences in the explicit acquisition of absolute pitch categories, it is important to highlight that the mechanisms of acquiring absolute pitch categories among a "true" AP population may be entirely different. For instance, if the phenomenon of AP was largely accounted for by individual differences in auditory working memory capacity, then one might expect AP to be much more common than the often cited statistic of one in every 10,000 individuals in Western cultures (e.g., Ward, 1999). That being said, an additional factor that might contribute to the rarity of "true" AP could be the prioritization of relative pitch information in early musical training, as relative pitch category learning has been shown to develop at a direct cost to absolute pitch sensitivity (e.g., Dye, Ramscar, & Suh, 2011). Nevertheless, the relationship we observed between auditory working memory, music experience, and the explicit learning of absolute pitch categories only accounted for approximately 39% of the variance in absolute pitch learning for Experiment 1, and approximately 24% of the variance in absolute pitch learning for Experiment 2. Accounting for this much variance in absolute pitch learning using just two variables (auditory working memory and age of music onset) is notable, though it is clear that there are likely several additional factors that influence the efficacy of learning absolute pitch categories (and these additional factors may be different for the acquisition of "true" AP ability).

These results point to a potential, unconventional strategy in learning absolute pitch categories. Given recent research suggesting that executive functions, such as working memory, can be improved through training (e.g., Harrison et al., 2013; Klingberg, 2010), it is likely that extended training on auditory working memory tasks – even if the working memory task does not specifically invoke memory for auditory pitch – would position an individual to better learn absolute pitch categories. Indeed, the fact that training on an auditory perceptual learning task has been shown to benefit general aspects of working memory suggests that perceptual learning – through intensively engaging working memory – shapes the general construction of working memory, apart from perceptual learning per se (Banai & Ahissar, 2009). In this line of reasoning, training to become more efficient at storing and manipulating perceptual events in working memory – even in a non-musical context – could very well sharpen the cognitive constructs necessary to successfully acquire absolute pitch categories.

The current studies provide an interesting perspective on the recent research examining mechanisms in perceptual category learning. According to the COVIS (competition between verbal and implicit systems) model of categorization (Ashby, Alfonso-Reese, Turken, & Waldron, 1998), working memory should offer a benefit to category learning tasks that favor a verbalizable, simple rule-based approach to learning. Categorization tasks that require a more implicit or procedural strategy might actually be negatively related to working memory ability (see DeCaro et al., 2008), since working memory might promote an inappropriate strategy for procedural or implicit category learning. The learning of absolute pitch categories does not fit with a verbalizable category learning framework (e.g., "high" versus "low" is verbalizable, but an inappropriate strategy because of octave equivalence), and thus the learning of absolute pitch categories seems like it should be thought of as an implicit category learning task. Yet, the way implicit categories are often tested is in an information-integration context (e.g., Ashby & Maddox, 2005), in which perceptual information must be integrated across more than one dimension in a non-verbalizable manner. In this sense, absolute pitch category learning does not fall within the realm of information-integration, as there is one salient perceptual dimension (pitch), and many orthogonal dimensions (e.g., loudness, harmonic spectrum, overall pitch height) that in theory are irrelevant for successful category identification.

Thus, the observed relationship between auditory working memory and absolute pitch category learning suggests the idea that working memory may not distinguish clearly between different types of category learning tasks or even different strategies in category learning (Lewandowsky et al., 2012). As the explicit acquisition of categories based on absolute pitch cues appears to be a particular case of category learning, not quite fitting into the mold of a verbalizable rule-based category structure or an implicit information-integration category structure, it was possible that we might have observed no relationship between category learning and working memory (or, perhaps, a negative relationship, as trying to explicitly use a verbalizable rule would have been an unsuccessful strategy). The fact that general auditory WM abilities predicted absolute pitch category learning in a non-AP population across two experiments suggests that participants with higher WM were able to selectively attend to one salient perceptual feature for categorization, while actively inhibiting other perceptual dimensions that might have inhibited learning. This view of working memory and category learning closely resembles what has been argued in speech, in which listeners must shift attention to diagnostic cues while filtering out information that is not diagnostic for learning (Nusbaum & Schwab, 1986). Indeed, working memory has been shown to predict the efficacy of understanding cochlear implant (CI) speech in both children (Fagan, Pisoni, Horn, & Dillon, 2007) and adults (Heydebrand, Hale, Potts, Gotter, & Skinner, 2007). CI speech is inherently a perceptual learning task, as individuals must learn to remap their acoustic-to-phonetic categories onto degraded speech sounds. We thus interpret the results from our current set of studies in broader (rather than a specifically musical) category-learning framework.

Our results should also be interpreted in the context of Gervain et al. (2013), who claim that adults taking the drug valproate learn absolute pitch categories significantly better than those taking a placebo, as valproate ostensibly reopens a critical period of learning. Importantly, the authors claim that participants who received valproate significantly improved in their absolute pitch category identification relative to placebo controls, correctly identifying an average of 5.09 out of 18 trials, or 28.3% post-training. This improvement (after a full week of training) is taken as evidence that absolute pitch – normally operating under a putative critical period of development – is “reopened” in individuals taking valproate, especially since the placebo control group did not show any reliable learning. Interestingly, the learning we demonstrate is the current set of studies is comparable to the observed learning in Gervain et al.’s initial valproate group, with individuals accurately classifying notes 31.0% of the time in the rote posttest, and 18.6% of the time in the generalization posttest (averaged across Experiments 1 and 2). Moreover, individuals who scored highly ($x > 1$ S.D.) on our working memory measures (INM in Experiment 1, auditory n-back in Experiment 2) were 43.8% accurate at identifying notes in the rote learning test, and 30.5% accurate at identifying notes in the generalization posttest. While it is true that the current set of studies was different in both the overall duration (one hour session vs multiple hours) and nature of training, the fact that high auditory WM individuals were performing in a comparable – if not superior – manner as the valproate participants in Gervain et al. (2013) suggests that perhaps a critical period terminology need not be applied.

Indeed, valproate has been suggested to indirectly affect dopamine (DA) release in the prefrontal cortex of rat brains (Ichikawa & Meltzer, 1999), which could have implications for working memory, as the phasic release of dopamine has been implicated in the updating of context information in the dorsolateral prefrontal cortex (D’Ardenne et al., 2012). Additionally, Yang, Lin, and Hensch (2012) found that the introduction of valproate significantly increases the cFos (an indirect measure of action potentials) in medial prefrontal cortex when adult mice were exposed to music, similarly suggesting that neuronal activity in frontal areas thought to be responsible for executive functions (e.g., working memory) differentially change in response to valproate and a sensory learning task. While these findings do not conclusively show that auditory working memory changes as a function of valproate, the current set of studies suggest that this alternative needs to be ruled out in order to make any definitive conclusions about reopening a critical period of learning.

In conclusion, across two experiments we found consistent evidence that an individual’s auditory working memory significantly predicted their subsequent performance on an absolute pitch category learning task, even when controlling for musical experience. These results suggest that the acquisition of intermediate absolute pitch ability (significantly above chance but below “true” AP performance) depends on an individual’s general auditory working memory ability, and that working memory ability mediates the relationship between musical factors (age of music onset) and absolute pitch learning. If this is the case, then it suggests that one reason why musical experience has been previously found to correlate with absolute pitch ability is because active musical instruction shapes domain-general auditory processes. Thus, while it may appear that specific musical factors are critical in determining the acquisition of absolute pitch categories, musical factors might only be important to the extent that they shape the general construction and use of auditory processes. In this sense, the explicit acquisition of absolute pitch categories in a non-AP population is perhaps best conceptualized as an exercise in perceptual category learning.

Acknowledgements

The authors would like to thank Gina Li for her help in data collection. This research was partly supported by an internal, University of Chicago funding source – the Norman H. Anderson Fund (to SCVH), as well as the Department of Defense through the Office of Naval Research under ONR N00014-13-1-0205 (to HCN).

References

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

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56*, 149–178.

Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., et al. (2007). Dichotomy and perceptual distortions in absolute pitch ability. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 104*(37), 14795–14800.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items, regardless of complexity. *Psychological Science, 18*(7), 622–628.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bachem, A. (1937). Various types of absolute pitch. *Journal of the Acoustical Society of America, 9*, 146–151.

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America, 27*, 1180–1185.

Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute pitch: An approach for identification of genetic and nongenetic components. *American Journal of Human Genetics, 62*, 224–231.

Bahr, N., Christensen, C. A., & Bahr, M. (2005). Diversity of accuracy profiles for absolute pitch recognition. *Psychology of Music, 33*, 58–93.

Banai, K., & Ahissar, M. (2009). Perceptual learning as a tool for boosting working memory among individuals with reading and learning disability. *Learning & Perception, 1*, 115–134.

Ben-Haim, M. S., Eitan, Z., & Chajut, E. (2014). Pitch memory and exposure effects. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 24–32.

Bermudez, P., & Zatorre, R. J. (2009). A distribution of absolute pitch ability as revealed by computerized testing. *Music Perception, 27*(2), 89–101.

Brady, P. T. (1970). Fixed-scale mechanism of absolute pitch. *Journal of the Acoustical Society of America, 48*, 883–887.

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

Chin, C. S. (2003). The development of absolute pitch: A theory concerning the roles of music training at an early developmental age and individual cognitive style. *Psychology of Music, 31*, 155–171.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*, 547–552.

Crozier, J. B. (1997). Absolute pitch: Practice makes perfect, the earlier the better. *Psychology of Music, 25*, 110–119.

Cuddy, L. L. (1968). Practice effects in the absolute judgment of pitch. *Journal of the Acoustical Society of America, 43*, 1069.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., & Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proceedings of the National Academy of Sciences USA, 109*, 19900–19909.

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.

Deutsch, D. (2002). The puzzle of absolute pitch. *Current Directions in Psychological Science, 11*, 200–204.

Deutsch, D. (2013). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 141–182). San Diego, CA: Academic Press.

Deutsch, D., & Dooley, K. (2013). Absolute pitch is associated with a large auditory digit span: A clue to its genesis. *Journal of the Acoustical Society of America, 133*, 1859–1861.

Deutsch, D., Henthorn, T., Marvin, W., & Xu, H. S. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for speech-related critical period. *Journal of the Acoustical Society of America, 119*, 719–722.

Deutsch, D., Le, J., Shen, J., & Li, X. (2011). Large-scale direct-test study reveals unexpected characteristics of absolute pitch. *Journal of the Acoustical Society of America, 130*, 2398.

Dye, M., Ramscar, M., & Suh, E. (2011). For the price of a song: How pitch category learning comes at a cost to absolute frequency representations. In *Proceedings of the 35th meeting of the cognitive science society*.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19–23.

Fagan, M. K., Pisoni, D. B., Horn, D. L., & Dillon, C. M. (2007). Neuropsychological correlates of vocabulary, reading, and working memory in deaf children with cochlear implants. *Journal of Deaf Studies and Deaf Education, 12*, 461–471.

Francis, A. L., & Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psychophysics, 71*, 1360–1374.

Gervain, J., Vines, B. W., Chen, L. M., Seo, R. J., Hensch, T. K., Werker, J. F., et al. (2013). Valproate reopens critical-period learning of absolute pitch. *Frontiers in Systems Neuroscience, 7*, 102.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review, 62*, 32–41.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585–612.

Gough, E. (1922). The effects of practice on judgments of absolute pitch. *Archives of Psychology (New York), 7*. i-93.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science, 24*, 2409–2419.

Hartman, E. B. (1954). The influence of practice and pitch distance between tones on the absolute identification of pitch. *American Journal of Psychology, 67*, 1–14.

Heald, S. L. M., Van Hedger, S. C., & Nusbaum, H. C. (2014). Auditory category knowledge in experts and novices. *Frontiers in Neuroscience, 8*, 260.

Hedger, S. C., Heald, S. M., & Nusbaum, H. C. (2013). Absolute pitch may not be so absolute. *Psychological Science, 24*, 1496–1502.

Heydebrand, G., Hale, S., Potts, L., Gotter, B., & Skinner, M. (2007). Cognitive predictors of improvements in adults' spoken word recognition six months after cochlear implant activation. *Audiology & Neurology, 12*, 254–264.

Ichikawa, J., & Meltzer, H. Y. (1999). Valproate and carbamazepine increase prefrontal dopamine release by 5-HT$_{1A}$ reception activation. *European Journal of Pharmacology, 380*, R1–R3.

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the *N*-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 615–622.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 333–358.

Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science, 25*, 334–339.

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology, 12*, 125–143.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?. Perception 36 ECVP Abstract Supplement 10.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14*, 317–324.

Kollmeier, B., Brand, T., & Meyer, B. (2008). Perception of speech and sound. In Jacob Benesty, M. Mohan Sondhi, & Yiteng Huang (Eds.), *Springer handbook of speech processing* (pp. 65). Springer.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Neuroscience, 11*, 599–605.

Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 830–845.

Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., et al. (2013). Resource allocation and prioritization in auditory working memory. *Cognitive Neuroscience, 4*, 12–20.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics, 56*, 414–423.

Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: Perception, coding, and controversies. *Trends in Cognitive Sciences, 9*, 26–33.

Lewandowsky, S. (2011). Working memory capacity and categorization: Individual differences and modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 3*, 720–738.

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

Lundin, R. W., & Allen, J. D. (1968). A technique for training perfect pitch. *Psychological Record, 12*, 139–146.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

Miyazaki, K. (1990). The speed of musical pitch identification by absolute pitch possessors. *Music Perception, 8*, 177–188.

Mull, H. K. (1925). The acquisition of absolute pitch. *American Journal of Psychology, 36*, 469–493.

Nusbaum, H. C., & Schwab, E. C. (1986). The role of attention and active processing in speech perception. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Volume 1, Speech Perception*. New York: Academic Press.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing, 30*, 653–661.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Preacher, K. J., & Hayes, A. F. (2008). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13–54). Thousand Oaks, CA: Sage.

Ross, D. A., Gore, J. C., & Marks, L. E. (2005). Absolute pitch: Music and beyond. *Epilepsy Behavior, 7*, 578–601.

Ross, D. A., Olson, I. R., & Gore, J. C. (2003). Absolute pitch does not depend on early musical training. *Annals of the New York Academy of Sciences, 999*, 522–526.

Rush, M. (1989). An experimental investigation of the effectiveness of training on absolute pitch in adult musicians. *Dissertation Abstracts International, 50*, 826A.

Russo, F. A., Windell, D. L., & Cuddy, L. L. (2003). Learning the "special note": Evidence for a critical period for absolute pitch acquisition. *Music Perception, 21*, 119–127.

Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology, 37*, 74–85.

Saffran, J. R., Reeck, K., Niebuhr, A., & Wilson, D. (2005). Changing the tune: The structure of the input affects infants' use of absolute and relative pitch. *Developmental Science, 8*, 1–7.

Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science, 14*, 262–266.

Schwab, E. C., Nusbaum, H. C., & Pisoni, D. B. (1985). Some effects of training on the perception of synthetic speech. *Human Factors, 27*, 395–408.

Sergeant, D. C. (1969). Experimental investigation of absolute pitch. *Journal of Research in Music Education, 17*, 135–143.

Smith, N. A., & Schmuckler, M. A. (2008). Dial A440 for absolute pitch: Absolute pitch memory by non-absolute pitch possessors. *Journal of the Acoustical Society of America, 123*, EL77–EL84.

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin, 113*, 345–361.

Takeuchi, A. H., & Hulse, S. H. (1991). Absolute pitch judgments of black-and white-key pitches. *Music Perception, 9*, 27–46.

Vianello, M. A. B., & Evans, S. H. (1968). Note on pitch discrimination learning. *Perceptual and Motor Skills, 26*, 576.

Ward, W. D., & Burns, E. M. (1982). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 431–451). San Diego, CA: Academic Press.

Ward, W. D. (1999). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 265–298). San Diego: Academic Press.

Wynn, V. T. (1973). Absolute pitch in humans: Its variations and possible connections with other known rhythmic phenomena. In G. A. Kerkut & J. W. Phillis (Eds.). *Progress in neurobiology* (Vol. 1, pp. 111–149). Elmsford, NY: Pergamon Press. Pt. 2.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440*, 91–95.

Yang, E. J., Lin, E. W., & Hensch, T. K. (2012). Critical period for acoustic preference in mice. *Proceedings of the National Academy of Sciences USA, 109*, 17213–17220.

Zatorre, R. J. (2003). Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience, 6*, 692–695.