A developmental study of latent absolute pitch memory

Kelly Jakubowski, Daniel Müllensiefen & Lauren Stewart

To cite this article: Kelly Jakubowski, Daniel Müllensiefen & Lauren Stewart (2016): A developmental study of latent absolute pitch memory, The Quarterly Journal of Experimental Psychology, DOI: 10.1080/17470218.2015.1131726

To link to this article: http://dx.doi.org/10.1080/17470218.2015.1131726

Published online: 15 Mar 2016.
A developmental study of latent absolute pitch memory

Kelly Jakubowski, Daniel Müllensiefen and Lauren Stewart

Department of Psychology, Goldsmiths, University of London, London, UK; Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark

ABSTRACT

The ability to recall the absolute pitch level of familiar music (latent absolute pitch memory) is widespread in adults, in contrast to the rare ability to label single pitches without a reference tone (overt absolute pitch memory). The present research investigated the developmental profile of latent absolute pitch (AP) memory and explored individual differences related to this ability. In two experiments, 288 children from 4 to 12 years of age performed significantly above chance at recognizing the absolute pitch level of familiar melodies. No age-related improvement or decline, nor effects of musical training, gender, or familiarity with the stimuli were found in regard to latent AP task performance. These findings suggest that latent AP memory is a stable ability that is developed from as early as age 4 and persists into adulthood.

ARTICLE HISTORY

Received 14 May 2015
Accepted 13 October 2015

KEYWORDS

Musical memory; Development; Absolute pitch

Memory for melodies involves encoding of both relational and absolute features of pitch. Most people remember melodies primarily in terms of relative pitch—that is, the relations or intervals between consecutive notes (Attneave & Olson, 1971). Melodies may also be remembered in terms of their absolute pitches, such that one could distinguish if a familiar piece of music was performed in a different key than usual (Terhardt & Seewann, 1983). A rare condition known as absolute pitch (AP) is the ability to label by name or produce single pitches without a reference tone. The prevalence of this ability is estimated at less than 1 in 10,000 people (Bachem, 1955; Profta & Bidder, 1988). However, aside from this rare ability (hereafter referred to as overt AP), a considerable amount of evidence has accumulated that a more latent form of absolute pitch memory is widespread in adults (e.g., Levitin, 1994; Schellenberg & Trehub, 2003; Terhardt & Seewann, 1983). This widespread ability has been referred to by multiple terms, including "implicit AP" (Deutsch, 2013; Frieler et al., 2013) and "residual AP" (Takeuchi & Hulse, 1993); the present paper uses the term "latent AP" (following Bartlette, Henry, & Moore, 2015; Jakubowski & Müllensiefen, 2013).

Latent AP is the ability to retain absolute pitch content from familiar melodies in the absence of explicit associations between single pitches and pitch names (Levitin, 1994). In adults, the ability to recall latent absolute pitch information appears to be independent of both formal musical training (Frieler et al., 2013; Levitin, 1994) and overt AP labelling ability (Jakubowski & Müllensiefen, 2013); however, latent AP has been found to correlate positively with relative pitch memory and certain emotional associations with melodies (Jakubowski & Müllensiefen, 2013). Latent AP information can also be retained by adults after only two hearings of a previously unfamiliar melody (Schellenberg, Stalinski, & Marks, 2014).

Although latent AP is widespread in adults, comparatively little is known about its developmental profile. Several theories on the acquisition of overt AP implicate a bias toward absolute processing of pitch at a young age, raising the question as to whether latent AP might also be enhanced early in life. These theories of overt AP development suggest
that all humans are born with the mechanisms for AP (e.g., Abraham, 1901) and generally implicate a “critical period” during which overt AP must be developed through training, typically before the age of 7 (e.g., Sergeant & Roche, 1973).

Evidence of absolute processing of pitch has been found in some infant studies, suggesting that the mechanisms involved may indeed be innate (Saffran & Griepentrog, 2001; Volkova, Trehub, & Schellenberg, 2006). For instance, Saffran and Griepentrog (2001) found that 8-month-old infants showed evidence of greater retention of absolute than relative pitch cues following exposure to unsegmented tone sequences. The authors suggest that this follows a general developmental trend from a bias towards perceiving absolute features as more salient toward more relational processing later in childhood, which is also seen in a variety of other domains (e.g., Gentner & Rattermann, 1991; Halford, 2005). Subsequent evidence for an absolute-to-relational shift was found by Stalinski and Schellenberg (2010). In their sample of 5- to 12-year-olds, children in all age groups were able to perceive differences in both absolute and relative pitch in a melody comparison task. However, younger children tended to rate transpositions of the same melody as a larger source of perceived difference between the two melodies than older children.

It is also well documented that children who begin musical training at a young age are much more likely to become overt AP possessors (Miyazaki, 1988, 1989). For example, Sergeant (1969) found that 87.5% of a group of musicians who began lessons at a mean age of 5.6 years possessed overt AP, and found no AP possessors in a comparison group who began lessons at a mean age of 9.9 years. Crozier (1997) attempted to teach identification of a single tone (A₄) to groups of young children (ages 4 to 5 years) and adolescents (ages 13 to 15 years). Following an initial training period, the children performed significantly more accurately in identifying the tone than the adolescents. Russo, Windell, and Cuddy (2003) conducted a similar study that compared three groups (3- and 4-year-olds, 5- and 6-year-olds, and adults) in terms of their ability to learn to identify the tone C₅. The group of 5- and 6-year-olds performed more accurately after training than the younger children and adults, providing evidence of a critical period around ages 5 to 6 years for AP development.

In terms of latent AP in children, similar accounts to theories on the aetiology of overt AP might be proposed, though evidence on this subject is sparse and not clear cut. Trehub, Schellenberg, and Nakata (2008) asked children to distinguish the correct version of a familiar song from a version that was pitch-shifted by one semitone. Canadian children from 5 to 8 years of age performed at chance level, and above-chance performance was found in 9- and 10-year-olds, although a subsequent analysis of variance (ANOVA) showed no overall effect of age. A second experiment found evidence for latent AP in 5- and 6-year-old Japanese children, leading the authors to suggest a cross-cultural processing difference between Western and Asian cultures. However, a subsequent study by two of the same authors (Schellenberg & Trehub, 2008) compared performance of Asian and non-Asian 9- to 12-year-olds on a latent AP task and found no significant differences between the two groups, although both groups performed significantly above chance level.

Several outstanding issues remain following these studies that the present work aimed to expand upon. Specifically, it is still not established as to whether non-Asian children under the age of 9 exhibit latent AP memory. This is crucial to understanding whether there exists a “critical period” for latent AP development similar to that of overt AP. As such, the present study measured latent AP memory in UK children across a wide age range (4 to 12 years) in order to capture any critical period that might exist for latent AP development. The present study also closely examined the role of familiarity with the musical stimuli used (TV theme tunes) on a trial-by-trial basis.

The present research had three main aims. The first was to test whether evidence of latent AP could be found in children and, if so, to determine its distribution. Some researchers have found evidence that overt AP is an “all-or-none”, bimodal ability (e.g., Athos et al., 2007), but evidence on the distribution of latent AP ability is sparse. Findings from Schellenberg and Trehub (2003, Figure 2) suggest that latent AP abilities exist on a continuum, rather than as a bimodal trait, in adults. The present research represents the first known exploration into this question

1However, see Trainor (2005) for an extensive critique of the stimuli and analysis used by Saffran and Griepentrog (2001) to compare relative versus absolute pitch judgments.
in children. Secondly, we aimed to investigate the developmental profile of latent AP to test whether (a) latent AP develops with age, similarly to various other musical abilities (Hannon & Trainor, 2007) and possibly in interaction with musical training, (b) latent AP is enhanced early in life and declines with age, as in theories of overt AP (e.g., Sergeant & Roche, 1973), or (c) latent AP is a stable trait exhibited throughout childhood and adolescence. Thirdly, we aimed to explore individual differences in performance potentially attributable to musical training, gender, and stimulus familiarity.

**Experiment 1**

**Method**

**Design**

We utilized a two-alternative forced-choice recognition task, similar to previous work on latent AP in children (Schellenberg & Trehub, 2008; Trehub et al., 2008) and adults (Schellenberg & Trehub, 2003). Participants heard two versions of an excerpt from a television theme tune—one version in the original key and one that was pitch-shifted (up or down) by one semitone—and were asked to choose which version sounded most like the one that they hear on TV at home.

**Participants**

The experiment was conducted as part of a residency at the Science Museum, London. The present paper only includes data from children who reported that they live in the United Kingdom, as the musical stimuli utilized reflected TV programmes aired regularly in the UK. We included for analysis only children who reported that they watched all five of their chosen TV programmes at least once per month, as some children retrospectively reported they had only seen a programme a couple times or had not seen the programme in several years. It is unlikely that such tunes would have stable long-term memory representations due to very infrequent exposure.

Participants who met these two prerequisite criteria were an opportunity sample of 200 children ages 4 to 12 years ($M = 7.98$, $SD = 2.22$; 115 female). Years of previous musical training ranged from 0 to 8 years ($M = 0.74$, $SD = 1.20$; see Table 1).

**Materials**

In order to provide a comprehensive stimulus set, a list of 59 television programmes was compiled. This list included programmes frequently watched by children in the UK (e.g., Charlie and Lola, Peppa Pig) and was based on TV viewing statistics and surveys distributed to parents and teachers of children in the target age range. Some popular programmes (e.g., Thomas the Tank Engine) were excluded on the basis of having had more than one theme song within recent years.

A 5-s sound clip was cut from the theme song for each of the 59 programmes. The clip was selected to represent the most recognizable section, generally from the chorus. Each clip was then pitched shifted twice using the software package Adobe Audition 3.0. The clip was shifted either up one semitone (ST) and then down two semitones (to create a $-1$ ST version) or down one semitone and up two semitones ($+1$ ST version). The original key version of the clip was also shifted twice (up one semitone, then down one semitone). This shifting method was utilized so that if artefacts were introduced into the stimuli by the shifting, the original key stimulus would be subjected to a similar amount of artefacts.

Four pairings of shifted and original key stimuli were created for each song (original/$+1$ ST, original/$-1$ ST, $+1$ ST/original, $-1$ ST/original) with 1 s of silence between each pair. One of these four orders was randomly chosen by the computer interface for each trial of each song. All sound clips were presented through Sennheiser RS 160 wireless headphones.

**Procedure**

Each child was tested individually in a sectioned-off area of the Science Museum, London. All participants had parental consent to take part.

Each participant was asked to look through the list of TV programmes and choose the five programmes he or she knew best. The list included the title of the TV programme along with a picture of the main characters to aid in identification. The experimenter explained that the experiment was a memory test,
so it was important to choose the TV programmes that the participant was most familiar with. The experimenter made and reviewed a list of five programmes with each child before testing commenced.

Instructions were read to each participant, explaining that the experimenter would play two short clips of music, separated by a brief silence. The participant would then be asked to indicate which version sounded most like the one he or she hears at home on TV. Participants were asked to make their best guess if they were unsure and were asked to close their eyes during each trial to help them to focus solely on the music.

The participant heard a practice trial, which was a song from a TV programme that he or she had not chosen for the actual experiment in order to become accustomed to the presentation format, length of the clips, and degree of pitch change between the clips.

The participant then proceeded to the actual experiment and answered a question after each trial regarding the frequency with which he or she watched that particular TV programme to assess relative familiarity (“How often do you generally watch this show?”: daily, a few times a week, weekly, monthly, less than monthly). At the end of the experiment, demographic information (age, gender, years of musical training, and nationality) was collected. The experiment lasted 5 to 10 min.

**Control study**

We conducted a control study analogous to those used in previous studies using a similar paradigm (Schellenberg & Trehub, 2003; Trehub et al., 2008), in which we aimed to ensure that the original key version of each TV tune was not able to be distinguished purely based on artefacts created during the pitch shifting process. This is particularly important as 44 of the 59 stimuli in our study contained vocals, which are more susceptible to distortion than instrumental stimuli (Trehub et al., 2008). In the control study, 10 adult participants completed the two-alternative forced-choice task. Each of the 59 TV tunes was presented once to each participant. As the purpose of the control study was to test whether the original version of a previously unfamiliar stimulus could be detected based purely on the sound quality, only the data for songs that the adult participants reported never having heard before were analysed. As a result, each participant contributed a total of 47 to 56 usable trials for analysis ($M = 53$); in total, 527 trials of the task were usable. Overall performance on this task was not significantly above chance level, $t(9) = 0.86$, $p = .41$ (mean accuracy = 51.8%, chance level = 50%). This indicates that the original key versions were not able to be distinguished purely based on pitch shifting artefacts.

**Results**

Figure 1 shows the distribution of scores across all participants. A total of 4.5% of participants answered correctly on all 5 trials, and 20% obtained a score of 4 out of 5 trials correct. The mean score was 2.71 out of 5, and 54.1% of all trials were answered correctly. To test the ability of the group of 200 children as a whole, a one-sample t test was employed to determine whether performance exceeded chance level of a score of 2.5. Overall, the group of children performed significantly above chance level, $t(199) = 2.60, p = .01$.

Next, we aimed to determine whether additional variables including age, gender, musical training, and familiarity ratings on a trial-level basis could help to explain individual differences in task performance (see Figure 2). A binomial mixed effects model was computed with the above-named variables as independent variables (fixed effects), participant included as a random effect, and correct/incorrect response on each trial as the dependent variable. The results of this analysis are presented in Table 2 and indicate no significant effect of any of the independent variables (all $ps > .05$). Trend analyses were also conducted to test specifically for an effect of age (the main predictor of interest to the present hypotheses) on task performance, and no significant linear, $F(1) =$
0.003, \( p = .96 \), or quadratic, \( F(1) = 0.26, p = .61 \), trends were revealed.

In order to further test for potential nonlinear age effects, a second binomial mixed effects model was computed where age was recoded as a categorical variable with the seven different age groups as factor levels. The previously non-significant fixed effects of gender, familiarity, and training were removed, and the present model included only age and participant as random effects. The fit of this model with age and participant as random effects was then compared against the fit of a model with only participant as a random effect, an intercept-only model without any random effect or predictor variables, and the full model with all predictor variables (as in Table 2) using \( \chi^2 \) tests. The results of this model comparison indicated that the four models did not significantly differ from one another (all \( ps > .05 \)). In addition, the Bayesian information criterion (BIC) indicated a better fit of the intercept-only model (BIC = 1386) than of the other three models (participant-only model: BIC = 1393; participant + age model: BIC = 1400; full model: BIC = 1426). This analysis suggests that the categorical variable age did not explain a significant proportion of the variance in the data, and it also indicates that there was no evidence for substantial individual differences between participants.

In addition to the above hypothesis tests, we ran an exploratory analysis in order to identify (a) potential nonlinear influences of any of the four predictor variables (age, gender, musical training, and familiarity), (b) potential effects arising from grouping adjacent age-years together (e.g., grouping 4- and 5-year-olds), and (c) potential higher order interaction effects among the four predictor variables. A conditional interference tree model was run that performs binary recursive partitioning on the dataset (Hothorn, Hornik, & Zeileis, 2006). The model did not identify any significant effects of any of the predictor variables or any interaction effects impacting on latent AP task performance.

The results of these analyses indicate that none of the predictor variables under investigation made significant contributions to explaining the variability
between participants. Hence we pooled all participants of our sample including all ages, musical backgrounds, and both genders in order to ask whether the difference in performance between participants is better understood by assuming one uniform group where all participants possess the same level of latent AP ability or by assuming two distinct groups where one group of participants might be performing at chance level but a second group would possess a higher ability level (i.e., whether latent AP ability is distributed on a continuum or bimodally). The latter would concur with the standard view of overt AP, which is commonly conceived as a binary ability that is present and stable in a small fraction of the general population and completely absent in the majority of people (Athos et al., 2007).

We tested this hypothesis by estimating two binomial mixture models with two versus one component, using maximum likelihood estimation via the Expectation-Maximisation algorithm as implemented in the R package FlexMix 2 (Grün & Leisch, 2008). The two-component model assigned 30 participants to the high-performing group and 170 participants to the low-performing group. However, according to the Bayesian information criterion the two-component model (BIC = 623) showed a considerably worse fit than the one-component model (BIC = 612), suggesting that participants’ performance in this sample is better modelled as arising from a uniform group with the same ability level. The estimated maximum-likelihood probability of recognizing a tune at the correct pitch level derived from the one-component model is .541 and coincides with the empirical success rate given above.

**Discussion**

In Experiment 1 we found overall above-chance performance on the latent AP task for the group of 200 children. No relationships were found between age, gender, musical training, or stimulus familiarity and task performance. Furthermore, from these data it seems that the above-chance ability to recognize tunes at the correct pitch level is a unitary trait that applies to all children in the sample equally. The present findings (mean accuracy = 54.1%) are similar to results found in studies of latent AP in adults (mean accuracy = 57.7%; Schellenberg & Trehub, 2003) using a one-semitone pitch shift comparison.

To further investigate latent AP ability in children, we designed a second experiment identical to Experiment 1 but utilizing two-semitone pitch-shifted stimuli. Several studies of adults have found significant improvements in task performance when the degree of shifting from the original pitch of the stimulus is increased (Schellenberg & Trehub, 2003; Smith & Schmuckler, 2008). Experiment 2 would thus allow us to ascertain the existence of latent AP at a second level of “resolution”. By decreasing the task difficulty we also aimed to further elucidate any individual differences contributing to latent AP in children.

**Experiment 2**

**Method**

**Design, materials, and procedure**

The materials and procedure employed were identical to those employed for Experiment 1 except that Experiment 2 utilized musical stimuli that had been shifted 2 semitones up or down from the original pitch to be distinguished from the original key stimuli.

**Participants**

Participants were 88 children aged 4 to 12 years (M = 7.75, SD = 2.39; 49 female). Years of previous musical training ranged from 0 to 5 years (M = 0.88, SD = 1.32; see Table 3).

**Control study**

Ten adults who were unfamiliar with the musical stimuli completed 503 trials of the task (42 to 58 trials each, M = 50 trials each). These participants did not perform significantly above chance, t(9) = 0.62, p = .55 (mean accuracy = 52.0%, chance level = 50%), suggesting that the original key version of each tune was not distinguishable purely based on pitch-shifting artefacts.

**Results**

Figure 3 shows the distribution of scores across all 88 participants. The mean score was 2.89 out of 5, and 57.7% of all trials were answered correctly. Overall,
the group of children performed significantly above
the chance level of a score of 2.5, \( t(87) = 3.59, p < .001 \).

An independent-samples t-test (adjusted for unequal variances) was conducted to compare the scores of the two groups of children in Experiments 1 and 2. The results indicate no significant difference in task performance between the two experiments, \( t(182.35) = -1.36, p = .18 \). A nonparametric Mann–Whitney U test on these data also gives a non-significant result, \( W = 8028.5, p = .22 \).

Next, we investigated individual differences due to age, gender, musical training, and familiarity (see Figure 4). As in Experiment 1, a binomial mixed effects model was computed, the results of which are presented in Table 4. Again, no significant effects of age, gender, musical training, or familiarity were found (all \( ps > .05 \)). Additionally, linear and quadratic trend analyses conducted to further examine any effects of age on task performance were non-significant [linear: \( F(1) = 0.12, p = .73 \), quadratic: \( F(1) = 0.04, p = .85 \)].

To further test for nonlinear effects of age, a model with age (as a categorical variable) and participant as a random effect, an intercept-only model, and the full model (as in Table 4) were compared using \( \chi^2 \) tests. The results of this model comparison revealed no significant differences between these four models (all \( ps > .05 \)). The Bayesian information criterion indicated a better fit of the intercept-only model (BIC = 605) than the model with only participant as a random effect (BIC = 611), the model that included both participant and age as random effects (BIC = 617), and the full model (BIC = 634). A conditional interference tree model was then built to test for nonlinear effects and higher order interactions of all predictor variables (age, gender, musical training, and familiarity). No significant effects were found for any of the predictor variables. In a final conditional interference tree analysis, the data from both Experiments 1 and 2 were pooled, due to the fact that the two datasets did not differ according to the previous \( \chi^2 \) tests, in order to maximize statistical power. Distance of pitch shifting of the stimulus (one or two semitones) was included as a factor. Again, no significant effects were found for any of the predictor variables on task performance.

Finally, we tested whether the performance scores on the latent AP task were better explained by a mixture of two participant groups with different ability levels or by a single group having the same ability level. For the data from Experiment 2, the two-component model again showed a worse fit to the data (BIC = 262) than the one-component model (BIC = 253). The estimated probability of recognizing a tune at the correct pitch level derived from the one-component model was .580.

**Discussion**

The results of Experiment 2 replicate those of Experiment 1 in all respects; the group as a whole performed significantly above chance on the task, no relationships were found between age, gender, musical training, or stimulus familiarity and task performance, and latent AP again appears to be a unitary rather than bimodally distributed trait. The absence of a significant difference between the data from Experiments 1 and 2 suggests that children perform similarly at differentiating stimuli that are shifted by one semitone or two semitones from the original version of a familiar tune.

**General discussion**

In this study, two separate groups of children aged 4 to 12 years performed significantly above chance in...
identifying the correct pitch level of familiar TV theme tunes. This represents the first known evidence of latent AP memory in children as young as 4 years. No evidence for linear or nonlinear effects of age on latent AP task performance was found in either experiment. This does not support the hypothesis that latent AP memory is enhanced early in life, as in theories of overt AP (Sergeant & Roche, 1973), nor the hypothesis that it develops steadily with age. Instead, latent AP appears to be a stable trait that is present from early childhood.

The origins and correlates of latent AP ability are still not well understood. None of the predictor variables employed in the present study (age, gender, musical training, or familiarity) made a substantial contribution to explaining the variance in the data. This parallels almost all other studies of latent AP in adults (e.g., Frieler et al., 2013; Levitin, 1994; Schellenberg & Trehub, 2003), which have also not identified any reliable predictors or correlates of latent AP. One exception is Jakubowski and Müllensiefen (2013), who found relative pitch memory and emotional associations with songs to be significant predictors of latent AP in adults. It is also yet to be determined whether latent AP relies on similar underlying perceptual or cognitive mechanisms to the rare ability of overt AP, as some authors have suggested these to be disparate abilities that require more distinct terminology (e.g., “absolute pitch” for the rare ability and “memory for key” or “memory for pitch level” for the more common ability; Schellenberg & Trehub, 2003; Trehub et al., 2008).

Contrary to the expectation that the level of stimulus familiarity might be a key influential factor, our results suggest that while children are better than chance level at recognizing the correct key for songs

Figure 4. Plots of age (in years), gender, musical training (in years), and familiarity rating (on a trial-level basis; 4 = most familiar) in relation to latent absolute pitch (AP) task scores (Experiment 2).

Table 4. Experiment 2: Mixed effects model of the influence of age, gender, musical training, and familiarity on latent AP task performance.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.47</td>
<td>0.46</td>
<td>1.04</td>
<td>.30</td>
</tr>
<tr>
<td>Familiarity</td>
<td>0.02</td>
<td>0.10</td>
<td>0.22</td>
<td>.82</td>
</tr>
<tr>
<td>Gender</td>
<td>0.21</td>
<td>0.20</td>
<td>1.07</td>
<td>.29</td>
</tr>
<tr>
<td>Age</td>
<td>−0.04</td>
<td>0.05</td>
<td>−0.79</td>
<td>.43</td>
</tr>
<tr>
<td>Musical training</td>
<td>−0.20</td>
<td>0.39</td>
<td>−0.52</td>
<td>.61</td>
</tr>
<tr>
<td>Age × Musical Training</td>
<td>0.02</td>
<td>0.04</td>
<td>0.63</td>
<td>.53</td>
</tr>
</tbody>
</table>

Note: AP = absolute pitch. Participant as random effect.
that are heard at least once per month, additional familiarity beyond this level (e.g., TV programmes watched once per day as compared to once per month) did not improve performance. One factor that could not be accounted for in the present design was cumulative exposure to a song across one’s lifetime. For instance, although an 8-year-old might currently only watch a TV programme once per week while a 4-year-old might watch the same programme once per day, the 8-year-old might have actually seen the TV programme more times in total during his lifetime due to the fact that he is twice as old as the 4-year-old. This cumulative exposure variable could potentially mask some age-related effects. However, the present design is not particularly suitable for addressing such an issue, as the question “How many times have you seen this TV programme in your life?” may be quite difficult for a young child to answer in a meaningful way. Future research should aim to further control the level of stimulus familiarity—for instance, by varying the number of hearings of previously unfamiliar melodies before testing latent AP memory—in order to determine the amount of cumulative exposure that is required for children of different ages to retain latent AP information.

Additionally, the data from the latent AP task within the present sample were more parsimoniously explained by a model that assumes latent AP to be a memory ability that exists on a continuum, rather than an “all-or-none”, bimodally distributed ability. This is in contrast to overt AP, which has often been referred to as a bimodal trait in adults (Athos et al., 2007; Baharloo et al., 1998), although some authors argue that overt AP may lie on more of a continuum than has previously been assumed (Takeuchi & Hulse, 1993; Vitouch, 2003). The present findings also parallel work from Schellenberg and Trehub (2003), which indicated that latent AP is a continuously distributed ability in adults.

Finally, we did not find a significant performance difference for one-semitone compared to two-semitone pitch-shifted stimuli (Experiment 1 versus Experiment 2). This is in contrast to previous findings in adults, in which performance on latent AP tasks improved significantly as distance of pitch shifting from the original key increased (Schellenberg & Trehub, 2003; Smith & Schmuckler, 2008). However, we did observe a trend toward better performance within Experiment 2 in line with previous findings, as the mean accuracy rate was 57.7% in Experiment 2 and 54.1% in Experiment 1.

A logical step for future research will be to explore this ability in infants to investigate whether latent AP memory is innate or whether it develops at some point before age 4. Some previous literature has suggested the absolute processing of pitch in infants as young as 6 months of age (Saffran & Griepentrog, 2001; Volkova et al., 2006), yet other studies have found no clear evidence of this ability in children of the same age (Plantinga & Trainor, 2005). These mixed results may be due to the variety of stimuli and paradigms previously employed. A paradigm similar to that used in the present study may produce more definitive results by utilizing highly naturalistic stimuli that children have been exposed to throughout their lives, rather than solely within the short confines of the experimental period.

In conclusion, the present study indicates that children as young as 4 years are able to distinguish the absolute pitch level of familiar melodies. Latent AP memory in children appears to be a widely distributed ability that is independent of age, gender, familiarity (over a prerequisite threshold of once a month exposure), and musical training.

Acknowledgements

We are grateful to Klaus Frieler, who designed the computer interface used for testing in both experiments. Thank you to Annick Odom and Michelle Dickinson who helped in testing participants. We thank the Science Museum, London for the opportunity to conduct this research within the museum as part of a Live Science residency in 2013.

References


