Measuring Children's Harmonic Knowledge with Implicit and Explicit Tests

Kathleen A. Corrigall, Barbara Tillmann, & E. Glenn Schellenberg

Measuring Children's Harmonic Knowledge with Implicit and Explicit Tests

Kathleen A. Corrigall[1,2], Barbara Tillmann[3,4], & E. Glenn Schellenberg[2,5]

[1]Department of Psychology, MacEwan University

[2]Department of Psychology, University of Toronto Mississauga

[3]CNRS, UMR5292; INSERM, U1028; Lyon Neuroscience Research Center,

Auditory Cognition and Psychoacoustics team, Lyon, F-69000, France

[4]University Lyon 1, Lyon, F-69000, France

[5]Centro de Investigação e Intervenção Social (CIS-IUL),

Instituto Universitário de Lisboa (ISCTE-IUL)

Author Note

Kathleen A. Corrigall https://orcid.org/0000-0003-2039-3932

Barbara Tillmann https://orcid.org/0000-0001-9676-5822

E. Glenn Schellenberg https://orcid.org/0000-0003-3681-6020

Funded by grants from the Social Sciences and Humanities Research Council of Canada awarded to KAC, and from the Natural Sciences and Engineering Research Council of Canada and the Fundação para a Ciência e a Tecnologia (FCT) em Portugal awarded to EGS. The Auditory Cognition and Psychoacoustics team of the Lyon Neuroscience Research Center is part of the LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-0060) of the Université de Lyon.

Correspondence concerning this article should be addressed to Katie Corrigall, Department of Psychology, MacEwan University, Edmonton, Alberta, Canada, T5J 4S2. E-mail: corrigallk@macewan.ca

Abstract

We used implicit and explicit tasks to measure knowledge of Western harmony in musically trained and untrained Canadian children. Younger children were 6-7 years of age; older children were 10-11. On each trial, participants heard a sequence of five piano chords. The first four chords established a major-key context. The final chord was the standard, expected tonic of the context or one of two deviant endings: the highly unexpected flat supertonic or the moderately unexpected subdominant. In the implicit task, children identified the timbre of the final chord (guitar or piano) as quickly as possible. Response times were faster for the tonic ending than for either deviant ending, but the magnitude of the priming effect was similar for the two deviants, and the effect did not vary as a function of age or music training. In the explicit task, children rated how good each chord sequence sounded. Ratings were highest for sequences with the tonic ending, intermediate for the subdominant, and lowest for the flat supertonic. Moreover, the difference between the tonic and deviant sequences was larger for older children with music training. Thus, the explicit task provided a more nuanced picture of musical knowledge than did the implicit task.

Keywords: music, harmony, children, development, music training

Measuring Children's Harmonic Knowledge with Implicit and Explicit Tests

Musicality (Honing, 2018) refers to musical knowledge that is a simple consequence of everyday exposure to music (e.g., Bigand & Poulin-Charronnat, 2006). One example is that individuals in Western societies with no music lessons have sophisticated expectations for pitch patterns in music (Schellenberg, 1996; Schellenberg et al., 2002; Tillmann et al., 2000), which, in turn, reveal an understanding of the hierarchical structure of pitch in Western music (Bigand et al., 1999, 2003; Krumhansl, 1990; Krumhansl & Cuddy, 2010). It can be a challenge to measure such knowledge, however, particularly among non-musicians and children, who cannot describe what they know using music-theoretical terminology. One alternative is to use implicit tasks that reveal musical knowledge indirectly. Implicit knowledge is acquired unintentionally and without awareness (Reber, 1967), and is difficult for a listener to describe. By contrast, explicit tasks require participants to make conscious judgments about the musical aspect of interest. For example, an explicit task might ask participants to rate how well a tone fits, continues, or completes a melody or chord sequence (e.g., Bigand & Pineau, 1997; Krumhansl, 1990; Schellenberg, 1996).

Implicit measures in the musical domain often use a priming paradigm (Bharucha & Stoeckig, 1986), asking listeners to make a simple binary judgment unrelated to the research question, such as identifying one of two instruments or sung syllables (e.g., Bigand et al., 2001; Schellenberg et al., 2005; Tillmann et al., 2006). The prime is the sequence of musical events that precedes the to-be-judged target event and typically establishes a musical key. The target (primed) stimulus is most often the final event of the sequence (but see Tillmann & Marmel, 2013). It is also the event that participants are asked to evaluate, and it varies in terms of how well it follows Western tonal rules. For example, if the prime implies the key of C major, the target could be an expected C major (tonic) chord, or a less expected chord (e.g., F major, subdominant). If listeners have the requisite knowledge of Western harmonic rules, their judgments should be faster when the final chord coincides with the expected tonic, which is easier to process, than when it coincides with a harmonically unexpected chord.

Although much research has been conducted using implicit (Bigand & Poulin-Charronnat, 2006) and explicit (Krumhansl, 1990) measures, less is known about what each task reveals about participants' musical knowledge, a question that motivated the present investigation with child participants. Implicit tasks appear to be extremely sensitive to measuring general tonal knowledge, such that response times are faster when the primed stimulus is the tonic rather than the subdominant (for review see Bigand & Poulin-Charronnat, 2006), even when explicit judgements fail to reveal harmonic knowledge (e.g., Tillmann et al., 2007; Marmel & Tillmann, 2009). Harmonic-priming effects are also evident in children (Schellenberg et al., 2005), high-functioning adolescents with Autism Spectrum Disorder (DePape et al., 2012), cerebellar patients (Tillmann et al., 2008), and individuals with acquired and congenital amusia (Tillmann et al., 2007, 2012; Omigie et al., 2012), although such effects are smaller in congenital amusics compared to controls. The harmonic-priming effect is cognitive rather than sensory because priming occurs when the tonic and subdominant have the same number of tones that overlap with the prime, and even when sensory overlap is greater for the less-expected subdominant chord (Bharucha & Stoeckig, 1987; Bigand et al., 2003; Marmel et al., 2010).

Because slower responses are thought to reflect additional processing of an event that violates expectations, one might expect larger harmonic-priming effects among individuals with more musical knowledge. After all, the degree to which a musical event is unexpected depends on the strength of an individual's expectations (i.e., tonal knowledge). Nevertheless, harmonic-priming effects rarely differ between groups that differ in musical knowledge. For example, studies of children reveal response-time differences that are independent of age, which coincides with increasing informal exposure to music, or of music training, which increases formal exposure (Marin, 2009; Schellenberg et al., 2005). Additional null results with music training have been reported in priming studies with adults (e.g., Bigand et al., 1999, 2001, 2003; Bigand & Pineau, 1997; Poulin-Charronnat et al., 2005; Tillmann et al., 1998; but see Loui & Wessel, 2007; Marmel & Tillmann, 2009). In short, harmonic-priming paradigms are particularly sensitive to measuring musical knowledge in general, but they appear to be fairly insensitive to individual differences in musical knowledge. Thus, it is possible that slower response times to unexpected sequences reflect a preconscious, earlier stage of processing that is not affected by the degree of harmonic violation because it does not yet include an evaluative component.

By contrast, studies that ask participants to make explicit evaluations of notes or chords, or to judge whether two musical sequences are the same or different, routinely report improvements as a function of age (e.g., Corrigall & Trainor, 2010, 2014; Cuddy & Badertscher, 1987; Kragness et al., 2021; Krumhansl & Keil, 1982; Lamont & Cross, 1994; Schellenberg et al., 2002; Trehub et al., 1986) or music training (e.g., Corrigall & Trainor, 2009; Marmel & Tillmann, 2009; Swaminathan et al., 2017, 2018; Wallentin et al., 2010). Explicit evaluations may reflect a later, conscious stage of processing that includes both a threshold effect—where enough tonal knowledge is required to differentiate between expected and unexpected sequences in the first place—as well as an evaluative component that indexes how strongly an unexpected sequence deviates from what is expected.

Based on these previous research findings, we predicted that our implicit task would not be as sensitive as our explicit task to the effects of musical experience acquired with age (i.e., accumulated listening over time) or music training (as measured by years of instrumental instruction). To test this prediction, we examined children from two age groups, who were musically trained or untrained in order to maximize differences in musical knowledge between participant groups. Based on previous research (see Corrigall & Schellenberg, 2015, for a review), younger children (6- to 7-year-olds) were expected to have only emerging behavioural (explicit) sensitivity to harmony, whereas 10- to 11-year-olds were expected to have accumulated more day-to-day exposure to music. Similarly, children enrolled in music lessons were assumed to have accumulated more formal musical knowledge.

Our implicit and explicit tasks used the same stimuli and were completed by the same participants, and both tasks measured listeners' knowledge of key membership and harmony in Western music. We compared children's responses to a standard, harmonically expected (tonic) ending to chord sequences with two levels of harmonic deviation: a highly unexpected (flat supertonic) violation to key membership, as well as a moderately unexpected (subdominant) harmonic deviation. We expected children to differentiate between standard and unexpected sequences in both the implicit and explicit tasks, but that effects of age, music training, and degree of violation would emerge only in the explicit task. Specifically, we predicted that older, musically trained children would differentiate most strongly between standard and deviant endings, particularly for highly unexpected key-membership violations.

Method

Participants

Listeners were typically developing 6- to 7-year-olds (n = 48) and 10- to 11-year-olds (n = 49) living in Canada (Table 1). Half of the children (n = 48) had no music training. The others had at least 9 months of lessons, but the older children (n = 27, M = 36.2, SD = 20.7) had significantly more lessons than the younger children (n = 22, M = 23.1, SD = 10.6), $t(40.35)$ = 2.87, p = .007, Cohen's d = .77 (unequal variances test). An additional 19 children were tested but excluded for inattention or missing data. A measure of socioeconomic status (SES) was formed by extracting the principal component from five measured variables (parents' education, family income, duration of the child's involvement in non-musical out-of-school activities, mother's and father's music training). As in previous research (e.g., Schellenberg, 2006; Swaminathan & Schellenberg, 2020), musically trained children ($M = 0.39$, $SD = 0.91$) came from higher-SES families than children without music training ($M = -0.39$, $SD = 0.94$), $t(95) = 4.15$, $p < .001$, $d = .92$.

Stimuli

Stimuli were eight sequences of five chords adapted from Corrigall and Trainor (2014). Each chord consisted of four notes in root position presented in piano timbre. Each sequence had three versions that were identical except for the last chord. The *standard* version ended on the tonic, the deviant *key* version ended on the flat supertonic, and the deviant *harmony* version ended on the subdominant (Figure 1). To avoid sensory priming (Bigand et al., 2003), neither the tonic nor subdominant chord occurred in the prime (the first four chords of each sequence), and tonic and subdominant targets shared the same number of pitch classes with the prime chords. The chord sequences were transposed to eight different major keys, from F to C Major in semitone increments. Each chord was 625 ms except for the final chord of each sequence, which was 1250 ms with an additional 1250-ms decay. Thus, sequences were approximately 5 s in duration.

For the implicit task, an additional version of each sequence was created, with the last chord presented in guitar timbre. Stimuli for training trials of the implicit task consisted of the same final chords in piano or guitar without the four-chord prime.

Procedure

Children were tested individually in a sound-attenuating booth. They wore headphones while sitting in front of an iMac computer. Each child was tested in two blocks of both the implicit and explicit tasks: a *key* block (tonic vs flat supertonic) and a *harmony* block (tonic vs subdominant). In both blocks, half of the trials had the standard, tonic-chord ending, whereas the other half had the deviant ending. The two implicit blocks always occurred first to allow for a completely naïve approach to the sequences, which would not be possible after the explicit task. Half of the participants were tested in the *key* block first and the *harmony* block second for both implicit *and* explicit tasks. The other half were tested in the reverse order.

For the implicit task, children completed 24 training trials (half piano, half guitar), followed by the two testing blocks (key, harmony), each with 64 test trials (8 sequences × 2 target chord endings × 2 timbres × 2 randomly selected major keys) presented in random order. Children were told that a cow named Milo was in a hurry and needed help sorting songs that ended with a piano or a guitar sound. They were then asked to respond as quickly as possible to single chords in the training phase, pressing one key for piano and another key for guitar. After the training and before the actual test phase that followed, children were told that they were now going to hear songs comprising five chords, and they should decide if the last chord was a guitar or piano. They then heard one example of a five-chord sequence that ended with piano, and one that ended with guitar. Feedback (correct, incorrect, too slow) was provided on all training and test trials. Although there was no time limit to respond, only correct responses that occurred within 2500 ms of the last chord were included in the statistical analyses. Excluded responses were uncommon (< 1%).

In the explicit task, all chords of the priming sequence and the target were presented with a piano timbre. Children were told that Milo now needed help deciding which songs to perform in a concert. They rated each song on a 5-point scale, with increasingly large ice cream cones corresponding to better-sounding songs. After confirming that they understood the rating scale, children completed two blocks of 32 trials (8 sequences × 2 target chord endings × 2 pseudo-randomly selected major keys) presented in random order. On each trial, children indicated their rating to the experimenter, who confirmed it with the child and entered the rating. There was no time limit for responding, and all responses were used to calculate children's scores.

After each of the four blocks, children took a 5-min break and played a word-search game. The entire procedure took approximately 1 hour, during which a parent completed a background questionnaire.

Results

Implicit task

In the implicit task, performance accuracy was high (> 90% correct) across age groups and conditions. Statistical analyses were restricted to response times (RTs; correct responses only) for piano trials,[1] which were averaged separately for each child, key and harmony blocks, and standard and deviant endings. Descriptive statistics are provided in Table 2.

RTs were analyzed with a mixed-design Analysis of Variance (ANOVA), which included block (key or harmony) and chordal ending (standard or deviant) as repeated measures, and age group (younger or older) and music training (trained or untrained) as between-subjects variables. There was a main effect of ending, $F(1, 93) = 28.82$, $p < .001$, partial $\eta^2 = .237$. RTs were faster for standard ($M = 1067$ ms, $SD = 224$) than for deviant ($M = 1121$ ms, $SD = 217$) endings (Figure 2, upper). There was also a main effect of age group, $F(1, 93) = 59.86$, $p < .001$, partial $\eta^2 = .392$. Older children ($M = 960$ ms, $SD = 156$) responded faster than younger children ($M = 1230$ ms, $SD = 179$) across conditions and endings. There were no main effects of block, $p = .082$, or music training, $p > .2$, and no two-way or higher-order interactions, *p*s > .1. Figure 3 (upper) illustrates difference scores (standard – deviant, absolute value) as a function of age group and music training.

[1] Results were identical when average RTs were calculated from piano and guitar trials combined (see Supplementary Materials).

The findings were identical when median rather than mean RTs were analyzed, with one exception (see Supplementary Materials). There was a small main effect of block, $F(1, 93) = 4.01$, $p = .048$, partial $\eta^2 = .041$. Median RTs were 26 ms faster in the key block than in the harmony block, but there were no interactions involving block, *p*s > .1.

Additional comparisons of RTs in response to piano and guitar target chords are provided in Supplementary Materials.

Explicit task

Ratings were averaged separately for each child for both blocks and both endings (Table 2). An ANOVA revealed a significant two-way interaction between block and ending, $F(1, 93) = 11.98$, $p < .001$, partial $\eta^2 = .114$ (Figure 2, lower). Although the difference between ratings for standard and deviant endings was significant in the key block, $F(1, 93) = 46.83$, $p < .001$, partial $\eta^2 = .335$, *and* in the harmony block, $F(1, 93) = 25.71$, $p < .001$, partial $\eta^2 = .217$, the interaction confirmed that it was larger in the key block. Additional follow-up analyses confirmed that ratings for the standard were similar across blocks, $F < 1$, whereas ratings for the deviant were lower in the key block, $F(1, 93) = 14.35$, $p < .001$, partial $\eta^2 = .134$. In other words, ratings were highest for standard tonic chords, lower for deviant subdominant chords, and lowest for deviant flat-supertonic chords.

There were no higher-order interactions involving block and ending, *p*s > .07, but there was a three-way interaction involving ending, age group, and music training, $F(1, 93) = 5.13$, $p = .026$, partial $\eta^2 = .052$. Difference scores (standard – deviant) are illustrated in Figure 3 (lower). For the younger children, the difference between standard and deviant endings was similar for those with or without music training, $F < 1$, and significant for both groups, *p*s < .05. For the older group, there was an interaction between ending and music training, $F(1, 47) = 7.46$, $p =$

.009, partial $\eta^2 = .137$. The difference between standard and deviant endings was significant for both groups, *p*s < .005, but stronger for children with music training.

We also examined whether response patterns might be affected by age-related differences in using the rating scale. Although a majority of children (72%) used the entire 5-point rating scale over the course of the testing session, a greater proportion of the older (19 of 49) compared to the younger (8 of 48) group used a relatively restricted range, $\chi^2(1, N = 97) = 5.90$, $p = .015$, $\varphi = .247$. Most of these children (23 of 27) simply avoided giving ratings of 1. When we transformed scores so that all children's data had the same range (1-5), response patterns were the same as in the original, unscaled analysis (see Supplementary Materials).

Individual differences

Global indices of harmonic knowledge were formed as difference scores (i.e., standard – deviant) separately for the implicit and explicit tasks. Because the two tasks were designed to measure the same construct, we expected a negative association between implicit and explicit difference scores, with deviant RTs longer than standard RTs for the implicit task, but standard ratings higher than deviant ratings for the explicit task. When collapsed across key and harmony blocks, however, scores in the implicit task were not correlated with scores in the explicit task, $r = -.085$, $N = 97$, $p > .4$. Moreover, no correlation between implicit and explicit measures was evident when performance was considered separately for the key, $r = -.158$, $N = 97$, $p > .1$, or the harmony, $r = -.032$, $N = 97$, $p > .7$, blocks. Thus, our measures of harmonic knowledge were statistically independent between the implicit and explicit tasks. In all instances, harmonic knowledge was independent of SES, *p*s > .4, and gender, *p*s > .1.

For the implicit task, scores in the key block were not correlated with scores in the harmony block, $r = .082$, $N = 97$, $p > .4$. For the explicit task, scores in the key block were correlated positively with scores in the harmony block, $r = .478$, $N = 97$, $p < .001$. Thus, within tasks, performance in the key and harmony blocks was correlated for the explicit, but not for the implicit measure.

We also considered further the results for the explicit task, which are illustrated in Figure 3. Perhaps the three-way interaction involving age, music training, and ending was a consequence of the fact that the older trained children had approximately 3 years of music lessons, whereas the younger trained children had only 2 years. We used multiple regression to model harmonic knowledge (difference scores) as a function of age and music training, treating both predictors as *continuous* variables. Because SES was associated with music training, it was also held constant. The model was significant, $R^2 = .125$, $F(3, 93) = 4.45$, $p = .006$, and both age, $\beta = .208$, $t(93) = 2.08$, $p = .041$, and music training, $\beta = .250$, $t(93) = 2.37$, $p = .020$, made significant independent contributions to the model. When the two-way interaction term (age X music training) was added to the model, it fell short of statistical significance, $p = .064$, although the slope was positive. In any event, the results confirmed that harmonic knowledge as measured by the explicit task increased in tandem with age and formal training in music.

Discussion

We used implicit and explicit tasks to examine knowledge of Western harmonic rules among children who varied in age and music training. The sample and the stimuli were identical across tasks—only the instructions and responses differed. Both tasks revealed knowledge of Western harmony.

In the implicit task, children exhibited harmonic priming, responding faster to the expected tonic than to moderately or highly unexpected chords. Response times to standards compared to deviants—an indication of harmonic sensitivity—were similar regardless of age, music training, or degree of harmonic violation. In the explicit task, children's ratings were highest for standards, intermediate for moderately unexpected deviants, and lowest for highly unexpected deviants, and the oldest, musically trained children made the largest distinction between standards and deviants. Thus, explicit ratings may be particularly sensitive to developmental trends, associations with music training, or differences between subtle versus obvious violations of musical rules.

Perhaps the difference between implicit and explicit measures stemmed from children's increased response confidence as age increased, which could have led them to choose more extreme scale levels. This hypothesis implies, however, that older children should have used a wider range of the rating scale, yet the opposite was true in our sample. Alternatively, our results could reflect differences in our dependent variables (response times compared to ratings), which may be difficult to dissociate from the underlying constructs that they are measuring. With child participants in particular, response times are likely to have more random variation (noise) compared to ratings, but less likely to be influenced by demand characteristics of the task (i.e., trying to provide the "correct" answer). Future research could attempt to disentangle potential influences of response uncertainty, response type, and musical knowledge by using a task with RT measures that correspond to an explicit assessment of harmonic knowledge (e.g., speeded judgements of whether the last chord belongs to the sequence or not, e.g., Schmuckler & Boltz, 1994).

The independence between our implicit and explicit tests of musical knowledge might be considered surprising because both tests were designed to measure the same construct. In two previous studies, adult amusic and control participants were compared on implicit *and* explicit tasks that measured expectancies for tones (i.e., melodic priming; Omigie et al., 2012, testing congenital amusics) or chords (i.e., harmonic priming; Tillmann et al., 2007, testing an acquired case of amusia). Both studies reported that the implicit task (response times) was a more sensitive measure of musical knowledge than was the explicit task (ratings). Indeed, both congenital and acquired amusics showed a priming effect with the implicit task, but a weaker effect (congenital amusics) or no effect (acquired amusic case) for the explicit task. Neither study reported correlations between tasks, however, and it would be interesting to further investigate this potential dissociation in the adult population (whether amusics or controls) in future research.

In studies of non-musical implicit knowledge that is learned in the laboratory, measures of implicit and explicit knowledge can also show different response patterns. For example, in artificial-grammar learning studies, participants are exposed initially to sequences of stimuli, with stimulus order constrained by the artificial grammar. When asked in the subsequent test phase whether sequences conform to the grammar, performance is above chance—indicating implicit learning—even though participants have no explicit knowledge of the grammatical rules (Reber, 1967). In studies of the mere-exposure effect (Bornstein, 1989), previous exposure to a stimulus leads participants to judge it more favorably in a subsequent test phase. In some instances, affective ratings for musical stimuli increase with exposure—indicating implicit memory, but recognition ratings (i.e., explicit memory) do not (Peretz et al., 1998; Szpunar et al., 2004). Explicit recognition is also more susceptible to stimulus manipulations (Peretz et al., 1998), whereas priming in general (Tulving & Schachter, 1990) and implicit learning in particular (Reber, 1992), are thought to be relatively inflexible across development and brain injury.

The dissociation observed here extends these findings to implicit harmonic knowledge that is acquired *outside* of the laboratory, through informal and/or formal exposure to Western music in everyday life, which could be another example of multiple memory systems, extended to conscious (explicit) and unconscious (implicit) knowledge of Western musical traditions. It is also consistent with the hypothesis that knowledge representations are graded, such that differences between implicit and explicit knowledge are "in degree rather than in kind" (Cleeremans & Jiménez, 2002; Destrebecqz & Cleeremans, 2002). From this perspective, the younger, musically untrained children relied on weaker, less stable representations that were less available to consciousness compared to the older trained children, who benefitted from more stable representations, which are necessary for explicitly graded judgments.

The distinction between implicit and explicit knowledge has been described previously using an *iceberg* metaphor, with the visible, above-water tip representing a relatively small part of the knowledge that an individual can articulate. The invisible, larger, below-surface part represents tacit or implicit knowledge, which is present but largely inaccessible, yet it informs attitudes, values, social norms, desires, and so on.[2] In the case of Western harmony, there is no doubt that evidence of this deeper knowledge can be captured with the harmonic-priming paradigm. The present findings suggest that the above-water tip represents knowledge influenced by conscious evaluation, which can, at least in some instances, provide a more nuanced picture of musicality.

[2] Although the metaphor is illustrative of Freud's theory and commonly attributed to Freud, "Freud never mentioned the iceberg in his published writings" (Green, 2019, p. 369).

References

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12*(4), 403-410. https://doi.org/10.1037//0096-1523.12.4.403

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: spreading activation or overlapping frequency spectra? *Perception & Psychophysics, 41*(6), 519-524. https://doi.org/10.3758/BF03210486

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 184-197. https://doi.org/10.1037/0096-1523.25.1.184

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception & Psychophysics, 59*(7), 1098-1107. https://doi.org/10.3758/BF03205524

Bigand, E., Poulin, B., Tillmann, B., Madurell, F., & D'Adamo, D. A. (2003). Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 159–171. https://doi.org/10.1037/0096-1523.29.1.159

Bigand, E., & Poulin-Charronnat, B. (2006). Are we ''experienced listeners''? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*(1), 100–130. https://doi.org/10.1016/j.cognition.2005.11.007

Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81*(1), B11-B20. https://doi.org/10.1016/S0010-0277(01)00117-2

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin, 106*(2), 265-289. https://doi.org/10.1037/0033-2909.106.2.265

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, philosophical and computational consensus in the making* (pp. 1-40). Psychology Press.

Corrigall, K. A., & Schellenberg, E. G. (2015). Music cognition in childhood. In G. MacPherson (Ed.), *The child as musician: A handbook of musical development* (2nd ed., pp. 415–429). Oxford, UK: Oxford University Press. https://10.1093/acprof:oso/9780198744443.003.0005

Corrigall, K. A., & Trainor, L. J. (2009). Effects of musical training on key and harmony perception. *Annals of the New York Academy of Sciences, 1169*, 164-168. https://doi.org/10.1111/j.1749-6632.2009.04769.x

Corrigall, K. A., & Trainor, L. J. (2010). Musical enculturation in preschool children: Acquisition of key and harmonic knowledge. *Music Perception, 28*(2), 195-200. https://doi.org/10.1525/mp.2010.28.2.195

Corrigall, K. A., & Trainor, L. J. (2014). Enculturation to musical pitch structure in young children: Evidence from behavioral and electrophysiological methods. *Developmental Science, 17*(1), 142-158. https://doi.org/10.1111/desc.12100

Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception & Psychophysics, 41*(6), 609–620. https://doi.org/10.3758/BF03210493

Destrebecqz, A. & Cleeremans, A. (2002). The self-organizing conundrum. *Behavioral and Brain Sciences, 25*(3), 334-335. https://doi.org/10.1017/S0140525X02270065

DePape A-M. R., Hall, G. B. C., Tillmann, B., & Trainor, L. J. (2012). Auditory processing in high-functioning adolescents with Autism Spectrum Disorder. *PLoS, ONE 7*(9): e44084. https://doi.org/10.1371/journal.pone.0044084

Green, C. D. (2019). Where did Freud's iceberg metaphor of mind come from? *History of Psychology, 22*(4), 369–372. https://doi.org/10.1037/hop0000135_b

Honing, H. (ed.) (2018). *The origins of musicality*. MIT Press.

Kragness, H. E., Swaminathan, S., Cirelli, L. K., & Schellenberg, E. G. (2021). Individual differences in musical ability are stable over time in childhood. *Developmental Science, 24*(4), e13081. https://doi.org/10.1111/desc.13081

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford University Press.

Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies in music. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 51–87). Springer. https://doi.org/10.1007/978-1-4419-6114-3_3

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory & Cognition, 10*(3), 243–251. https://doi.org/10.3758/BF03197636

Lamont, A., & Cross, I. (1994). Children's cognitive representations of musical pitch. *Music Perception, 12*(1), 27-55. https://doi.org/10.2307/40285754

Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception & Psychophysics 69*(7), 1084–1092. https://doi.org/10.3758/BF03193946

Marin, M. M. (2009). Effects of early musical training on musical and linguistic syntactic abilities. *Annals of the New York Academy of Sciences, 1169*, 187-190. https://doi.org/10.1111/j.1749-6632.2009.04777.x

Marmel, F., & Tillmann, B. (2009). Tonal priming beyond tonics. *Music Perception, 26,* 211-221. https://doi.org/10.1525/mp.2009.26.3.211

Marmel, F., Tillmann, B., & Delbé, C. (2010). Priming in melody perception: Tracking down the strength of cognitive expectations. *Journal of Experimental Psychology: Human Perception & Performance, 36*(4), 1016-1028. https://doi.org/10.1037/a0018735

Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia, 50*(7), 1483–1493. https://doi.org/10.1016/j.neuropsychologia.2012.02.034

Peretz, I., Gaudreau, D., & Bonnel, A. M. (1998). Exposure effects on music preferences and recognition. *Memory & Cognition, 15*, 379 –388. https://doi.org/10.3758/BF03201171

Poulin-Charronnat, B., Bigand, E., & Madurell, F. (2005). The influence of voice leading on harmonic priming. *Music Perception, 22*(4), 613-627. https://doi.org/10.1525/mp.2005.22.4.613

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behaviour, 6*(6), 855–863. https://doi.org/10.1016/S0022-5371(67)80149-X

Reber, A. S. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition, 1*(2), 93–113. https://doi.org/10.1016/1053-8100(92)90051-B

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition, 58*(1), 75-125. https://doi.org/10.1016/0010-0277(95)00665-6

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology, 98*(2), 457–468. https://doi.org/10.1037/0022-0663.98.2.457

Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General, 131*(4), 511–537. https://doi.org/10.1037/0096-3445.131.4.511

Schellenberg, E. G., Bigand, E., Poulin, B., Garnier, C., & Stevens. C. (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science, 8*(6), 551-566. https://doi.org/10.1111/j.1467-7687.2005.00447.x

Swaminathan, S., Schellenberg, E. G., & Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence, 62*, 119–124. https://doi.org/10.1016/j.intell.2017.03.005

Swaminathan, S., Schellenberg, E. G., & Venkatesan, K. (2018). Explaining the association between music training and reading in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(6), 992–999. https://doi.org/10.1037/xlm0000493

Swaminathan, S., & Schellenberg, E. G. (2020). Musical ability, music training, and language ability in childhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(12), 2340-2348. http://dx.doi.org/10.1037/xlm0000798

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 370-381. https://doi.org/10.1037/0278-7393.30.2.370

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*(4), 885-913. https://doi.org/10.1037/0033-295X.107.4.885

Tillmann, B., Bigand, E., Escoffier, N., & Lalitte, P. (2006). The influence of musical relatedness on timbre discrimination. *European Journal of Cognitive Psychology, 18,* 343–358. https://doi.org/10.1080/09541440500269548

Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research, 61*(3), 157-174. https://doi.org/10.1007/s004260050022

Tillmann, B., Gosselin, N., Bigand, E., & Peretz, I. (2012). Priming paradigm reveals harmonic structure processing in congenital amusia. *Cortex, 48*(8), 1073-1078. https://doi.org/10.1016/j.cortex.2012.01.001

Tillmann, B., Justus, T., & Bigand, E. (2008). Cerebellar patients demonstrate preserved implicit knowledge of association strengths in musical sequences. *Brain and Cognition, 66*(2), 161-167. https://doi.org/10.1016/j.bandc.2007.07.005

Tillmann, B., & Marmel, F. (2013). Musical expectations within chord sequences: Facilitation due to tonal stability without closure effects. *Psychomusicology, 23*(1), 1-5. https://doi.org/10.1037/a0030454

Tillmann, B., Peretz, I., Bigand, E., & Gosselin, N. (2007). Harmonic priming in an amusic patient: The power of implicit tasks. *Cognitive Neuropsychology, 24*(6), 603–622. https://doi.org/10.1080/02643290701609527

Trehub, S. E., Cohen, A. J., Thorpe, L. A., & Morrongiello, B. A. (1986). Development of the perception of musical relations: Semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 295-301. https://doi.org/10.1037/0096-1523.12.3.295

Tulving, E. & Schachter, D. L. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306. https://doi.org/10.1126/science.2296719

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test: A new reliable test for measuring musical competence. *Learning and Individual Differences, 20*(3), 188–196. http://dx.doi.org/10.1016/j.lindif.2010.02.004

Table 1

Means (and SDs) of Demographic Variables as a Function of Age Group and Music Training (G = Girls; B = Boys).

	Younger		Older	
	Untrained	Trained	Untrained	Trained
	n = 26	*n* = 22	*n* = 22	*n* = 27
Age in Years	7.07 (.46)	7.05 (.39)	10.96 (.41)	10.92 (.40)
Music Training in Months	0 (0)	23.1 (10.6)	0 (0)	36.2 (20.7)
Socioeconomic Status	-.317 (1.06)	.344 (.768)	-.485 (.798)	.419 (1.02)
Gender (G/B)	13/13	15/7	8/14	13/14

Note: Socioeconomic status was the principal component extracted from five measured variables: family income (*n* = 86), parents' education (*n* = 93), duration of child's involvement in non-musical out-of-school activities (*n* = 93), mother's music training (*n* = 88), and father's music training (*n* = 87), such that 0 was the mean for the entire sample (*SD* = 1). Missing data on the measured variables were substituted with the mean.

Table 2.

Means (and SDs) as a Function of Task, Testing Condition, Standard or Deviant Ending, Age Group, and Music Training. Response Times for the Implicit Task are in Milliseconds. Ratings for the Explicit Task are on a Scale that Ranged from 1 (sounds bad) to 5 (sounds good).

	Condition			
	Key		Harmony	
	Standard	Deviant	Standard	Deviant
Implicit Task (Response Times)				
Younger—Untrained	1201 (258)	1236 (214)	1244 (260)	1304 (256)
Younger—Trained	1176 (166)	1235 (164)	1194 (179)	1239 (173)
Older—Untrained	942 (170)	981 (173)	981 (205)	1031 (234)
Older—Trained	900 (151)	983 (151)	915 (172)	967 (156)
Explicit Task (Ratings)				
Younger—Untrained	3.69 (.81)	3.19 (.81)	3.74 (.71)	3.54 (.64)
Younger—Trained	4.06 (.69)	3.66 (.67)	3.95 (.62)	3.71 (.62)
Older—Untrained	3.42 (.46)	2.98 (.68)	3.54 (.48)	3.25 (.45)
Older—Trained	3.95 (.61)	2.66 (.89)	3.81 (.65)	3.19 (.80)

Figure 1. An example of the stimulus chord sequences. The first four chords (A minor, E minor, B minor, D major) establish the key of G major. The fifth chord is the standard tonic ending (G major, upper), the flat supertonic (A-flat major, middle), or the subdominant deviant ending (C major, lower).

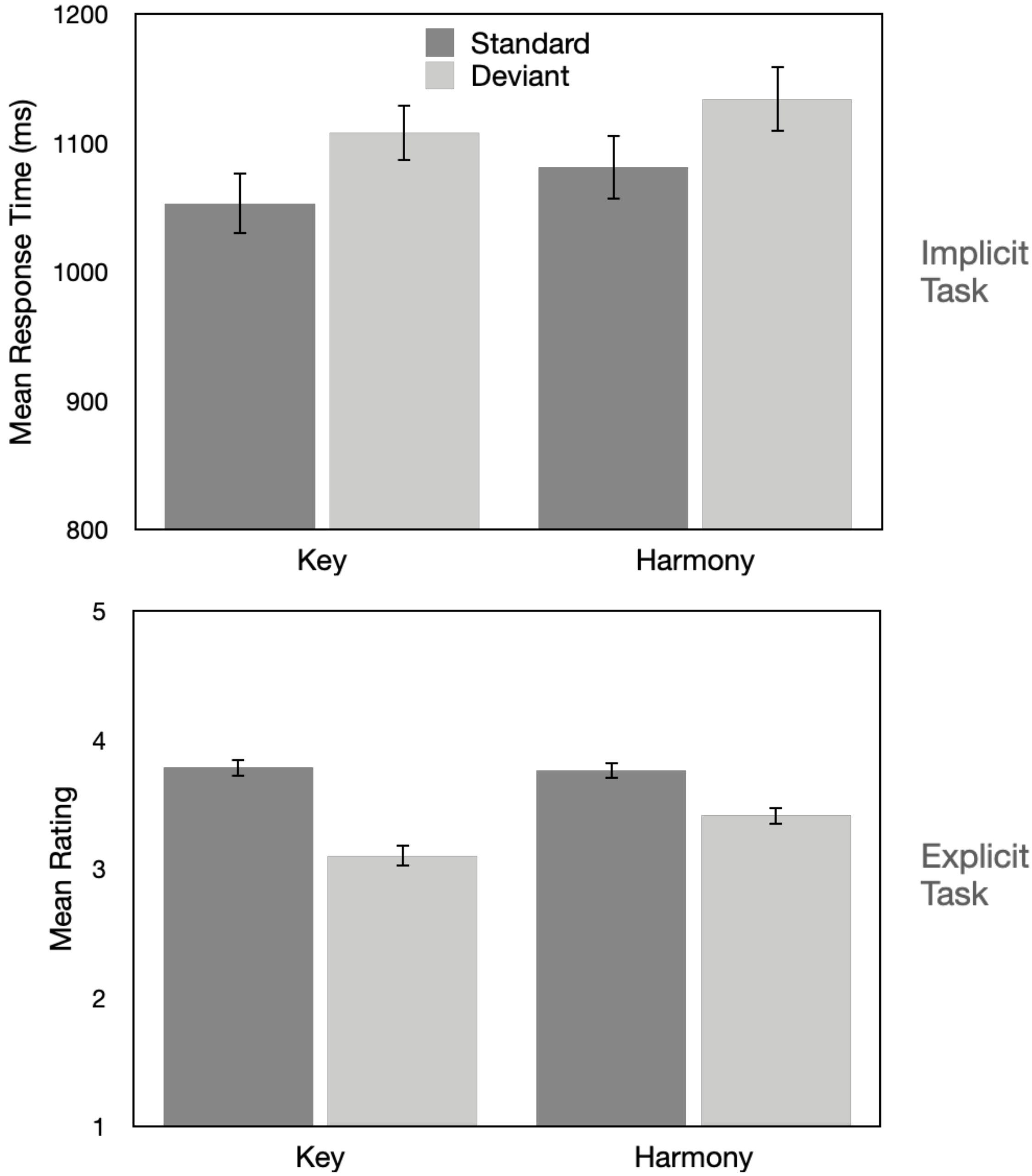

Figure 2. Response patterns from the implicit (upper) and explicit (lower) tasks as a function of condition (key or harmony) and chordal ending (standard or deviant). For the implicit task, there was no interaction between condition and ending. For the explicit task, ratings were higher for standard than for deviant endings, and the difference was greater in the key condition than in the harmony condition. Error bars are standard errors of the mean.

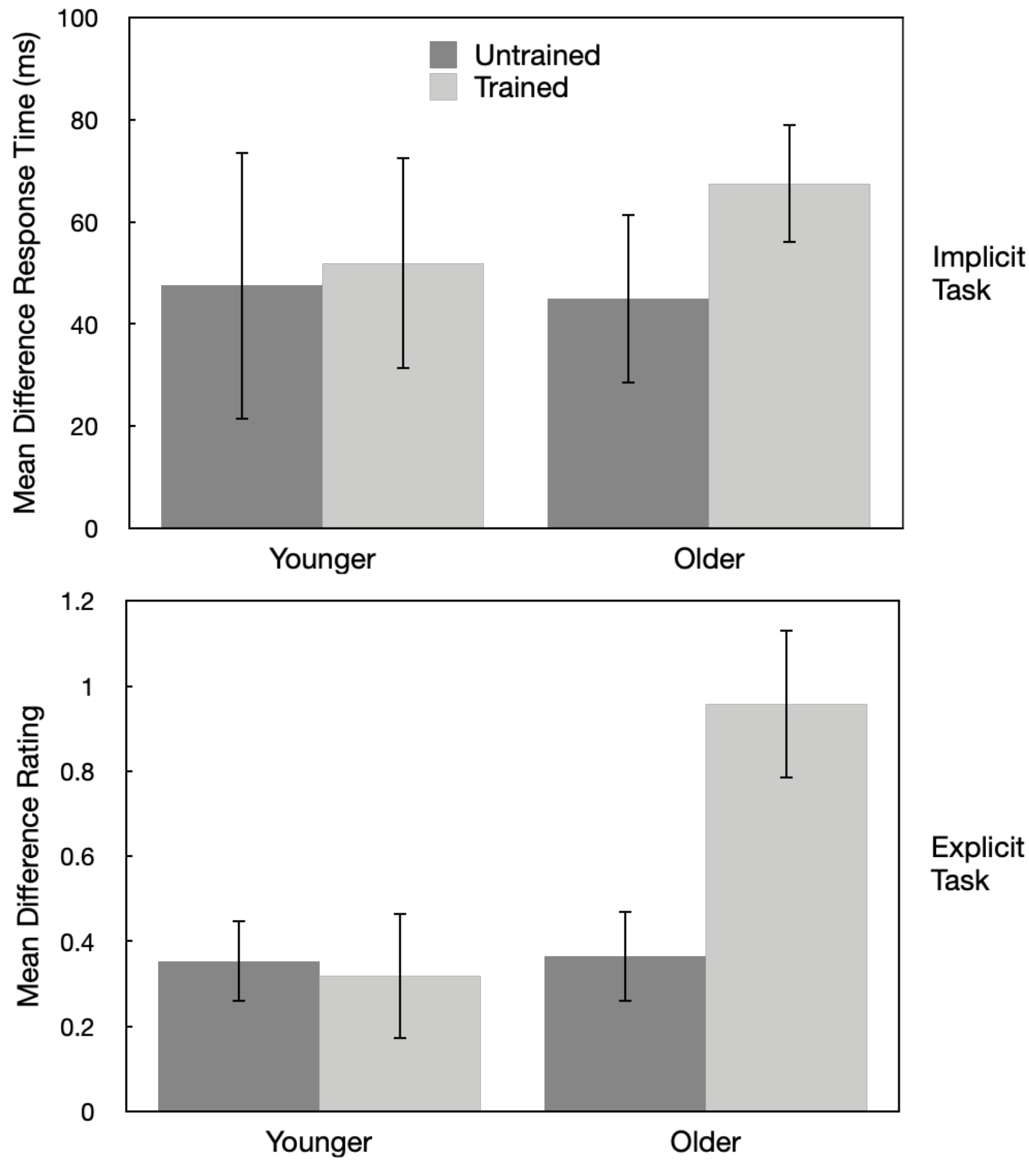

Figure 3. Difference scores (standard endings – deviant endings) collapsed across key and harmony conditions, for response times in the implicit task (upper) and ratings in the explicit task (lower). For the implicit task, absolute values are illustrated. In both panels, the difference between standard and deviant sequences in absolute terms was greater for older children with music training compared to other children, but the three-way interaction (ending × age group × music training) was significant only for the explicit task. Error bars are standard errors of the mean.

Supplementary Materials

"Measuring Children's Harmonic Knowledge with Implicit and Explicit Tests"

Corrigall, Tillmann, & Schellenberg

In the main manuscript, we reported results from the implicit task, specifically response times (RTs) averaged across correct responses to piano trials because these were the same stimuli that were used in the explicit task. Here, we continue to consider only correct responses, but we provide three additional analyses: (1) *median* RTs in response to piano trials, (2) mean RTs *combined* across piano and guitar trials, and (3) mean RTs considered *separately* for piano and guitar trials. We also provide an additional analysis of data from the explicit task, in which we transformed some of the children's data. Specifically, scores from children who failed to use the entire rating scale were transformed so that all children's data ranged on a scale from 1 to 5.

Implicit Task, Median Response Times—Piano Trials

We re-ran the ANOVA reported in the main manuscript, substituting median RTs for mean RTs. In other words, for each child, we analyzed the median RT identified separately for standard and deviant endings for the harmony blocks of trials, and for standard and deviant endings for the key blocks. This analysis ensured that the results from the original ANOVA were not the unintended consequence of long RTs for some children, which could have had undue influence on mean RTs.

As in the original analysis, we conducted a four-way ANOVA with two repeated measures: ending (standard or deviant) and block (harmony or key), and two between-subjects variables: age group (younger or older) and music training (trained or untrained). There were again significant main effects of ending, $F(1, 93) = 30.12$, $p < .001$, partial $\eta^2 = .245$, and age, $F(1, 93) = 57.00$ $p < .001$, partial $\eta^2 = .380$. Median RTs were 68 ms faster for standard than for deviant endings, and 262 ms faster for the older group compared to the younger group of children. There was also a small but statistically significant main effect of block, $F(1, 93) = 4.01$, $p = .048$, partial $\eta^2 = .041$. Response times, averaged across standard *and* deviant endings, were 36 ms faster in the key block compared to the harmony block.

Implicit Task, Mean Response Times Averaged Across Piano and Guitar Trials

The next analysis considered mean RTs averaged across piano *and* guitar trials. As in the original analysis, a four-way ANOVA with two repeated measures (ending, block) and two between-subjects variables (age group, music training), revealed a main effect of ending, $F(1, 93) = 18.75$, $p < .001$, partial $\eta^2 = .168$, and a main effect of age, $F(1, 93) = 72.38$, $p < .001$, partial $\eta^2 = .438$. Mean RTs were 24 ms faster for standard compared to deviant endings, and 247 ms faster for older compared to younger children.

Implicit Task, Mean Response Times Analyzed Separately for Piano and Guitar Trials

The next analysis considered timbre as an additional independent variable, such that a five-way ANOVA had three repeated measures (ending, block, timbre) and two between-subjects variables (age group, music training). Main effects of ending, $F(1, 93) = 20.70$, $p < .001$, partial $\eta^2 = .182$, and age, $F(1, 93) = 72.94$, $p < .001$, partial $\eta^2 = .440$, remained significant, with faster responding for standard compared to deviant endings (by 25 ms), and for older compared to younger children (by 250 ms). A main effect of timbre was also evident, $F(1, 93) = 220.11$, $p < .001$, partial $\eta^2 = .703$, with more rapid identification of timbre when the target chord contrasted with that of the priming sequence, such that guitar targets were identified 163 ms faster than piano targets.

Both main effects of ending and timbre were qualified, however, by a two-way interaction between ending and timbre, $F(1, 93) = 21.29$, $p < .001$, partial $\eta^2 = .186$. The ending manipulation was significant for piano targets, $F(1, 93) = 28.82$, $p < .001$, partial $\eta^2 = .237$, with a 53-ms advantage for standard compared to deviant endings (i.e., the principal finding in the main manuscript). For guitar targets, the ending manipulation had no effect, $F < 1$, and the RT difference was only 2 ms. In short, the main effect of faster responding for standard than for deviant endings was due entirely to piano trials, which also explains why the ending effect was weaker in the previous analysis, when RTs were averaged across timbres. This finding is in also consistent with previous musical-priming studies, which show stronger priming effects for targets that are: (1) in-tune or consonant compared to out-of-tune or dissonant, (2) presented in the same or a similar timbre as the context compared to those presented in a different or dissimilar timbre (see Tillmann et al., 2006, for a discussion).

One additional significant result from the five-way ANOVA was a small two-way interaction between block and timbre, $F(1, 93) = 6.65$, $p = .011$, partial $\eta^2 = .067$. In the harmony block, the RT advantage for guitar trials was 181 ms (averaged across standard and deviant endings), $F(1, 93) = 144.51$, $p < .001$, partial $\eta^2 = .608$. In the key block, the RT advantage for guitar trials was only 144 ms, yet stronger in terms of the effect size, $F(1, 93) = 179.13$, $p < .001$, partial $\eta^2 = .658$. This result was unexpected and difficult to interpret. Presumably, the salience of the timbre change, which was of no theoretical interest, influenced the salience of the other manipulations.

Explicit Task, Ratings Transformed so that All Children's Scores Were on the Same Scale

For the 27 children who failed to use the entire range, 23 children did not assign ratings of 1, two did not assign ratings of 5, one did not assign ratings of 1 or 5, and one did not assign ratings of 4 or 5. We transformed the scores from these children so that they all came from a scale in which the lowest rating was 1 and the highest rating was 5. For example, for the 23 children who did not assign ratings of 1, scores were transformed according to the following formula (X = original rating, Y = transformed rating):

$$Y = X - 1 + (X - 2) / 3$$

This formula ensured that an original score of 2 was transformed to 1, a score of 3 was transformed to 2.33, and scores of 4 and 5 were transformed to 3.67 and 5, respectively.

The results of the mixed-design ANOVA remained unchanged. There was a significant two-way interaction between block and ending, $F(1, 93) = 13.27$, $p < .001$, partial $\eta^2 = .125$. The interaction confirmed that the difference between ratings for standard and deviant endings was larger in the key block, $F(1, 93) = 50.66$, $p < .001$, partial $\eta^2 = .353$, compared to the harmony block, $F(1, 93) = 25.10$, $p < .001$, partial $\eta^2 = .213$. Ratings for the standard were similar across blocks, $F < 1$, whereas ratings for the deviant were lower in the key block, $F(1, 93) = 14.58$, $p < .001$, partial $\eta^2 = .136$. As in the original analysis, then, ratings were highest for standard tonic chords, lower for deviant subdominant chords, and lowest for deviant flat-supertonic chords.

The three-way interaction involving ending, age group, and music training remained significant, $F(1, 93) = 5.92$, $p = .017$, partial $\eta^2 = .060$. For the younger children, the difference between standard and deviant endings was similar for those with or without music training, $F < 1$, and significant for both groups, *p*s < .05. For the older group, there was an interaction between ending and music training, $F(1, 47) = 9.08$, $p = .004$, partial $\eta^2 = .161$. The difference between standard and deviant endings was significant for both groups, *p*s < .003, but stronger for children with music training.

Additional Reference

Tillmann, B., Bigand, E., Escoffier, N., & Lalitte, P. (2006). The influence of musical relatedness on timbre discrimination. *European Journal of Cognitive Psychology, 18*(3), 343–358. https://doi.org/10.1080/09541440500269548