

Widespread but Limited in Context: Absolute Tuning Judgments are Disrupted by Relative Pitch Cues

Stephen C. Van Hedger^{1,2,3} and Noah R. Bongiovanni⁴

Abstract

Most listeners exhibit absolute pitch memory for familiar pitched sounds, ranging from simple tones (e.g., the “bleep” used to censor broadcast media) to rich, polyphonic melodies (e.g., excerpts from popular songs). Given that relative pitch is the predominant means of processing music for most listeners, this observed absolute pitch memory suggests that most listeners have hybrid representations, containing both absolute and relative pitch cues. The present experiment assesses how relative pitch cues influence absolute pitch memory in a variety of listening contexts, varying in terms of ecological validity. Participants were asked to explicitly judge different musical sounds (isolated notes, chords, scales, and short excerpts), varying in complexity, in terms of absolute tuning (i.e., whether a musical sound adhered to conventional, A440 tuning). Listeners showed robust above-chance performance in judging absolute tuning when the to-be-judged sounds did not contain relative pitch changes (i.e., isolated notes and chords), replicating prior work. Critically, however, performance was significantly reduced and was not statistically different from chance when the musical sounds contained relative pitch changes (scales and short excerpts). Taken together, the results suggest that most listeners possess some long-term memory for absolute tuning, but this ability has limited generalizability to more ecologically valid musical contexts that contain relative pitch changes.

Keywords

Absolute pitch, memory, perception, relative pitch, music cognition

Submission date: 16 May 2023; Acceptance date: 19 September 2023

When one listens to music, what features of the experience are retained in memory? The answer to this question depends on several factors, including whether an individual possesses absolute pitch (AP). AP is generally defined as the ability to name or produce any musical note without the aid of a reference note (e.g., Deutsch, 2013; Levitin & Rogers, 2005; Loui, 2016; Takeuchi & Hulse, 1993) and is an extremely rare ability, with a base rate of approximately 0.01% of the global population (Bachem, 1955). AP possessors, in contrast to the general population, are able to form explicit categories based on pitch *chroma*, also called pitch class. Importantly, these categories are rooted in the ability to identify the fundamental frequency of the note; as such, other characteristics of a musical note, such as its timbre, octave register, or relative pitch context in which the note is heard, do not negate the AP possessor’s ability to correctly identify the pitch by its

note name (Takeuchi & Hulse, 1993). For example, the authors, both having AP themselves, could recognize and identify that a “G” played on a solo piano is the same “G” when produced by a French horn during a symphony

¹ Department of Psychology, Huron University College at Western, London, ON, Canada

² Department of Psychology, University of Western Ontario, London, ON, Canada

³ Western Institute for Neuroscience, University of Western Ontario, London, ON, Canada

⁴ Department of Music, University of Notre Dame, Notre Dame, USA

Corresponding author:

Stephen C. Van Hedger, Department of Psychology, Huron University College at Western, 1349 Western Road, London, ON N6G 1H3, Canada.

Email: svanhedg@uwo.ca



played with a full orchestra; the note retains its “G” *chroma* whether heard alone or in concert, and irrespective of other cues.

In contrast, the majority of listeners process music using relative pitch (RP), meaning that notes are not remembered by their *chroma*, but instead by how each note relates to the other notes heard in musical contexts (e.g., Levitin & Rogers, 2005; Miyazaki, 1995; Plantinga & Trainor, 2005; Saffran et al., 2005; Takeuchi & Hulse, 1993). RP processing, as a result, is thought to lead to unstable representations of pitch chroma. For example, under an RP framework, a “G” heard in the context of a musical work in G major will sound “congruent” and is likely to receive a higher goodness-of-fit rating than, say, a “G” heard in the context of C-sharp major, where the pitch “G” is “incongruent” and would be likely to receive a lower goodness-of-fit rating (see Krumhansl, 1979). Subsequent work has found that this tonal hierarchy also influences memory processes, with congruent notes being more likely to result in false memories than incongruent notes (Vuvan et al., 2014).

From this perspective, AP and RP can be conceptualized as opposed constructs – two non-overlapping means of processing pitch that, in some instances, might interfere with one another (Ziv & Radin, 2014). Yet, an accumulating body of research suggests that all listeners – regardless of whether they possess AP – can perceive and remember musical pitch in both absolute and relative terms, which suggests that pitch processing is complex and might depend on the listening context (e.g., Creel & Tumlin, 2012; Hoeschele, 2017). Despite not possessing explicit labels based on pitch chroma, which characterizes the phenomenon of AP, several converging findings have indicated that most listeners have remarkably good absolute pitch memory. For example, non-AP possessors are able to identify the correct-key version of a familiar piece of music from other versions shifted by as little as one semitone (e.g., Schellenberg & Trehub, 2003; Terhardt & Seewann, 1983; Van Hedger et al., 2018a), even in children as young as four (Jakubowski et al., 2017). This general finding is by no means limited to musical sounds; several experiments have demonstrated similarly high performance among non-AP listeners asked to identify the correct-pitch versions of non-musical sounds, including the North American landline dial tone (Smith & Schmuckler, 2008) and the 1,000 Hz censor “bleep” used in broadcast media (Van Hedger et al., 2018b). These demonstrations of absolute pitch memory for non-musical pitches are particularly intriguing when compared to AP, as they reflect good pitch memory for a single pitch class (conceptually similar to traditional tests of AP, which require the identification of isolated pitch classes).

Yet, it is worth noting that the absolute pitch memory abilities surveyed in the previous paragraph necessarily rely on prior knowledge of specific pitched sounds. In other words, there is something specific (e.g., an iconic recording of a popular song) that individuals recall in

making their judgments. Although listeners have preserved absolute pitch memory for novel instrumental versions of familiar recordings (Van Hedger et al., 2023), suggesting a form of generalization, these judgments nevertheless require at least some prior familiarity with the iconic recording. The critical importance of familiarity is highlighted by findings that have convincingly demonstrated that listeners perform at chance when making similar judgments for unfamiliar recordings of music (e.g., Schellenberg & Trehub, 2003). In contrast, an AP possessor is not thought to be limited by familiar context; a well-known and an unknown recording both possess an easily identifiable chroma or key signature to someone with AP. This observation highlights one of the critical presumed differences between AP and non-AP possessors with respect to pitch memory and additionally raises the question of whether this widespread absolute pitch memory, observed in non-AP possessors, generalizes to broader listening contexts.

An emerging body of research has suggested that widespread pitch memory may be generalizable beyond specific, highly familiar sounds. Although these findings are still predicated on familiarity, they are presumably based on a more abstract familiarity with specific pitch classes and tuning systems, rather than specific familiarity with a given iconic sound, such as a familiar recording. For example, non-AP listeners’ judgments of preference, tonal fit, and tension depends on the familiarity of the pitch class and musical key, as measured by frequency of occurrence in Western music (e.g., C versus F#; Ben Haim et al., 2014; Eitan et al., 2017). In another study, non-AP possessors were able to judge whether isolated notes were tuned conventionally (where the “A” above middle C is set to 440 Hz – hereafter referred to as A440 tuning) or unconventionally (shifted from the conventional standard by a quarter tone), once again presumably based on the familiarity of hearing these sounds (Van Hedger et al., 2017). This sensitivity to the absolute tuning standard of A440 among non-AP possessors provides a parallel to the phenomenon of AP, which is also a form of absolute hearing. These studies demonstrate a further degree of abstraction, as it is unlikely that participants are basing their judgments on any iconic recording of isolated notes or simple chord progressions. These results, taken together, suggest that listeners’ memories for absolute pitch reflect more generalized knowledge (i.e., extending beyond specific iconic recordings) about how frequently particular pitch classes are encountered or whether a specific pitch would count as a “real” (correctly tuned) note in Western music’s specific cultural tuning standard.

Thus, the current state of the literature on absolute pitch memory reflects tensions along the dimensions of both generalizability and ecological validity. On the one hand, listeners display robust absolute pitch memory for highly familiar and ecologically rich sounds, such as popular music recordings (e.g., Schellenberg & Trehub, 2003). This absolute pitch memory generalizes to novel versions

of the familiar melody (Van Hedger et al., 2023) but nevertheless is inherently tied to prior experience with the iconic recording, which suggests a limited implicit recognition of pitch chroma beyond these familiar sounds. On the other hand, recent work (Ben-Haim et al., 2014; Van Hedger et al., 2017) has demonstrated much more abstract memories based on pitch chroma (e.g., through testing isolated notes), ostensibly based on broader distributional regularities of hearing particular pitch classes and tunings. However, this more abstracted form of absolute pitch memory has hitherto suffered from relatively low ecological validity, as it has been tested using isolated notes specifically lacking relative pitch cues, which may interfere with absolute pitch processing (cf. Ziv & Radin, 2014). Understanding the extent to which absolute and relative pitch cues interact in pitch memory assessments is critical in characterizing the similarities and differences in how both AP and non-AP possessors represent musical pitch.

The present experiment was therefore designed to assess how absolute pitch judgments for novel sounds would interact with the availability of relative pitch cues. In the present experiment, participants judged whether musical sounds adhered to conventional (A440) tuning across several types of musical stimuli differing in complexity and ecological validity (isolated notes, chords, scales, and excerpts from musical compositions). Given that individuals were recruited regardless of their formal musical training, the tuning judgment task was framed to participants in terms of valence (with conventionally tuned items sounding “good” and unconventionally tuned items sounding “bad”), as has been done in earlier work (Van Hedger et al., 2017). Grounding our hypothesis in the prior literature on pitch memory among non-AP possessors, we predicted that listeners would be above chance in differentiating conventionally tuned from unconventionally tuned isolated musical notes, conceptually replicating prior work (Van Hedger et al., 2017). However, given the importance of relative pitch processing in music (e.g., Miyazaki et al., 2018) and the discussion of how relative and absolute pitch processing might interfere with one another (Miyazaki, 1993, 1995), we additionally hypothesized that tuning judgments for stimuli with salient relative pitch cues would be attenuated (i.e., it would be harder to differentiate conventionally tuned from unconventionally tuned sounds) relative to simpler stimuli such as isolated notes. Overall, this experiment was designed to test whether listeners demonstrated evidence of long-term memory for absolute tuning when using more complex, ecologically valid musical sounds in which RP cues are increasingly present.

Method

Participants

A total of 200 participants were recruited from Amazon Mechanical Turk, and 193 were retained for the final

analyses ($M = 40.43$ years, $SD = 12.03$ years, range of 21 to 70 years). The *Data Exclusion* section provides details on participant exclusion. We used Cloud Research (Litman et al., 2017) to recruit a subset of participants from the larger Mechanical Turk participant pool. Participants were required to have at least a 90% approval rating from a minimum of 50 prior Mechanical Turk assignments and had to have passed internal attention checks administered by Cloud Research. All participants received monetary compensation upon completion of the experiment.

Materials

The experiment was programmed in jsPsych 6 (de Leeuw, 2015). The to-be-judged musical sounds were all created in (or imported into, in the case of the musical excerpts) Reason 4 (Propellerhead: Stockholm) and used a piano timbre. Given that the musical sounds were initially generated in MIDI format, the tuning of each sound was altered prior to exporting each as an audio file. Specifically, the in-tune sounds were exported by setting the master tuning option to the default (+0 cents; A440 tuning), whereas the out-of-tune sounds were exported by setting the master tuning option to +50 cents (~A453 tuning). All sounds had a sampling rate of 44.1 kHz and 16-bit depth.

The isolated musical notes were similar to those reported by Van Hedger et al. (2017). We tested 12 notes, ranging from C4 to B4. Each of the 12 notes in this range had 2 versions – an in-tune version (e.g., C4) and an out-of-tune version (e.g., C4 + 50 cents) – resulting in 24 total stimuli. The isolated musical chords consisted of 12 major triads in root position, with the root note ranging from C4 to B4. Like the isolated notes, each of the 12 major chords in this range had two versions – an in-tune version (e.g., C4 major) and an out-of-tune version (e.g., C4 major + 50 cents) – resulting in 24 total stimuli. The musical scales consisted of the first 5 notes of 12 major scales, with the initial note of each scale ranging from C4 to B4. Like the isolated notes and chords, each of the 12 major scales in this range had two versions – an in-tune version (e.g., C4 major scale) and an out-of-tune version (e.g., C4 major scale + 50 cents) – resulting in 24 total stimuli. The musical notes, chords, and scales were all 1,000 ms in duration. The musical excerpts were selected from four multi-movement piano compositions. Nine excerpts were selected from *Opus 109* (18 Etudes) by Friedrich Burgmüller (1858), seven excerpts were selected from *Petite Suite* by Alexander Borodin (1885), six excerpts were selected from *Opus 165* (España) by Isaac Albéniz (1890), and eight excerpts were selected from *Suite Española* by Isaac Albéniz (1886). These particular pieces were selected because each contained several short piano movements and were hypothesized to be generally unfamiliar to participants. Although there were 30 total musical excerpts, only 24 were randomly selected to be presented to each participant to keep the design similar to the

other Conditions. Half of the randomly selected excerpts were tuned conventionally (A440) and half were tuned unconventionally (~A453). Each excerpt was trimmed to 5,000 ms, with a 500 ms linear fade in and fade out.

Inter-trial “scramble” sounds were designed to disrupt echoic and short-term memory of the preceding musical stimulus (cf. Nees, 2016). Each inter-trial scramble sound consisted of 1/*f* pink noise, which was mixed with *tone clouds*. The tone clouds contained 50 sine tones per second – in other words, a new tone onset occurred every 20 ms. Each sine tone was 75 ms in duration, and the tones were randomly generated within a range of 147–1,320 Hz. The scramble sounds were 5,000 ms in duration, with the tone clouds spanning over the first 4,500 ms of the sound (i.e., 225 tones). The tone clouds and pink noise were created with a custom script in Matlab (MathWorks: Natick, MA).

The auditory calibration stimuli consisted of a 30-s brown noise, generated in Adobe Audition (Adobe: San Jose, CA), as well as sine tones adapted from Woods et al. (2017), which were designed to assess participant headphone use. All sounds presented in the experiment were root-mean-square normalized to the same level (−20 dB relative to full spectrum).

Procedure

After reading a letter of information outlining the details of the study and providing informed consent, participants completed the auditory calibration portion of the experiment. This consisted of (1) listening to the 30-s calibration noise and adjusting their device’s volume to a self-determined comfortable listening level and (2) completing a headphone assessment first reported by Woods et al. (2017), in which participants listened to three tones on each trial and made a forced-choice judgment as to which tone was quieter than the other two. There were six total judgments in the headphone assessment, and a score of five or six correct responses was taken as evidence that the participant was wearing headphones. Following the auditory calibration, participants were randomly assigned to one of four Conditions: (1) Notes, (2) Chords, (3) Scales, or (4) Excerpts. Regardless of Condition, participants were instructed that they would be making judgments about the tuning of piano sounds. Specifically, participants were told that some of the sounds would be *in tune* and some would be *out of tune*. In case participants had little or no musical training, *in tune* was also defined as sounding *familiar* or *good*, whereas *out of tune* was also defined as sounding *unfamiliar* or *bad*. Participants were additionally instructed to base their decisions on their initial, “gut” feeling.

Participants then completed the main task, in which they judged 24 musical sounds (12 conventionally tuned, 12 unconventionally tuned) in a randomized order. Each trial consisted of (1) the sound presentation, (2) a forced-choice judgment of whether the sound was *in tune* or *out of tune*,

and (3) a 5,000 ms inter-trial “scramble” sound to disrupt immediate echoic memory and minimize inter-trial carry-over effects. Participants repeated this procedure until all 24 musical sounds had been rated. At the end of the main task, participants were presented with two auditory attention checks, in which they were instructed to click on a specifically marked button on the computer via spoken instructions.

Following the main rating task, participants completed a short questionnaire. The questionnaire assessed age, gender, highest level of education, hearing aid use, primary and secondary languages, self-reported musical ability, self-reported pitch perception, formal musical training, childhood musical appreciation, and childhood musical encouragement. Participants were also given a free response option to disclose any information they thought might influence their results. Participants were then provided with a unique completion code and compensated.

Data Exclusion

There were three primary reasons why participants were excluded from analysis. The first was providing an incorrect answer for the auditory attention checks administered in the experiment. Given that these attention checks occurred without prior warning, failing to provide a correct answer was taken to indicate that a participant had not consistently listened to the musical stimuli. The second reason was if participants reported the use of a hearing aid or otherwise indicated that they had a health concern that might affect the results of the study. The third was if participants self-reported possessing absolute pitch, which was a response option on the self-reported pitch perception question. Headphone use was not used to exclude participants, as headphone use was recommended but not required. Overall, 126 participants (63%) passed the headphone screening measure.

Of the 200 recruited participants, one failed to correctly answer both auditory attention assessments, one participant reported the use of a hearing aid, one participant disclosed that their perception of pitch had significantly deteriorated with age, and four participants reported possessing absolute pitch. This left 193 participants in the primary analysis (Note Condition: $n = 45$, Chord Condition: $n = 51$, Scale Condition: $n = 52$, Excerpt Condition: $n = 45$).

Data Analysis

To assess whether participants in each Condition were above chance in determining whether the musical sounds were conventionally tuned versus unconventionally tuned, we used one-sample *t*-tests, with mean accuracy being tested against the chance estimate of 50%. To assess whether mean accuracy differed across Conditions, we used a one-way analysis of variance (ANOVA) with Condition (Notes, Chords, Scales, Excerpts) as a between-participant factor. Reported post-hoc tests used Bonferroni–

Holm corrections. Correlations between questionnaire measures and performance were assessed via uncorrected Pearson product-moment correlations. Finally, for each analysis we report both p -values and Bayes factors – calculated via Bayesian equivalents of the reported tests in JASP (Marsman & Wagenmakers, 2017) – to facilitate the understanding of the relative evidence for or against our hypothesis. The reported Bayes factors (BF_{10}) represent how likely the alternative hypothesis is, relative to the null hypothesis, given the data. For the p -values, we adopted a significance threshold of $\alpha = .05$. Statistics are reported according to the APA Publication Manual (7th Edition), with exact p -values being reported for values greater than .001.

Results

Mean Performance Tested against Chance

Overall, we found clear evidence for above-chance performance when the to-be-judged musical sounds did not contain temporally unfolding relative pitch cues (Figure 1). Mean performance in the isolated Note Condition was 57.4% (SD : 11.0%), which was significantly above chance, $t(44) = 4.51$, $p < .001$, $d = 0.67$; $BF_{10} = 463.2$. Mean performance in the isolated Chord Condition was 57.8% (SD : 11.1%), which was also significantly above chance, $t(50) = 5.17$, $p < .001$, $d = 0.73$; $BF_{10} = 4,405$.

In contrast, the musical scales and musical excerpts – which both contained relative pitch changes that unfolded over time – did not show clear above-chance performance. Mean performance in the musical Scale Condition was 52.1% (SD : 9.5%), which did not statistically differ from the chance estimate, $t(51) = 1.58$, $p = .120$, $d = 0.22$, $BF_{10} = 0.484$. Mean performance in the musical Excerpt Condition was nominally lower, at 51.2% (SD : 9.9%), which did not statistically differ from chance, $t(44) = 0.82$, $p = .418$, $d = 0.12$, $BF_{10} = 0.221$. Notably, the Bayes factor from the musical excerpt analysis suggests that the null hypothesis (i.e., at-chance performance) is approximately 4.5 times more likely than the alternative hypothesis given the data.

Comparisons across Conditions

The one-way ANOVA found a significant main effect of Condition, $F(3,189) = 5.56$, $p = .001$, $\eta_p^2 = .081$, $BF_{10} = 22.56$, suggesting that at least one group significantly differed from one or more of the other groups. Post-hoc tests supported the hypothesis that the isolated notes and chords differed from the sounds containing relative pitch changes (scales and excerpts). Isolated musical notes were significantly higher than both scales, $t = 2.52$, $p = .038$; $BF_{10} = 3.67$, and excerpts, $t = 2.83$, $p = .025$; $BF_{10} = 6.58$. Isolated musical chords were significantly higher than both scales, $t = 2.90$, $p = .021$; $BF_{10} = 8.49$, and excerpts, $t = 3.21$, $p = .009$; $BF_{10} = 15.75$. Isolated

notes and chords did not significantly differ from one another, $t = -0.28$, $p = 1$; $BF_{10} = 0.22$, nor did scales and excerpts significantly differ from one another, $t = 0.42$, $p = 1$; $BF_{10} = 0.23$.

Correlating Performance with Questionnaire Measures

The results of the exploratory¹ correlational analyses are reported in Table 1. None of the measures collected in the questionnaire were associated with tuning performance accuracy. The only significant correlations we observed were between age and self-reported music skill, as well as strong intercorrelations between our music-related questionnaire measures (self-reported musical skill, self-reported pitch perception, the extent to which music listening was encouraged during childhood, the extent to which music listening was appreciated during childhood, and whether the participant reported formal musical training). The lack of an association between tuning performance accuracy and our collected measures was not driven by a particular Condition – that is, running the correlations separately for each Condition yielded comparable (i.e., non-significant) results.

Discussion

The present experiment assessed whether listeners could accurately judge the absolute tuning of different kinds of musical sounds varying in ecological richness. Although prior work has suggested that absolute memory for tuning is widespread (Van Hedger et al., 2017), this prior work only tested isolated musical notes, which have little ecological validity. Thus, in making claims that absolute tuning might influence the aesthetic evaluation of music (e.g., Farnsworth, 1968) or even health-related outcomes (Calamassi & Pomponi, 2019), it is important to determine under what circumstances listeners can make these judgments. The results of this experiment suggest that listeners can make accurate judgments of absolute tuning under limited circumstances. Specifically, listeners appear to be robustly above chance in determining whether “static” notes and chords, which do not contain relative pitch changes, are conventionally (versus unconventionally) tuned. Strikingly, this ability is significantly attenuated (and does not significantly differ from chance) as soon as the to-be-judged sounds contain relative pitch changes. In this sense, the results provide a direct replication of Van Hedger et al. (2017) but also suggest that absolute tuning perception is only observed in highly artificial settings, in which the to-be-judged sounds do not contain relative pitch changes. Overall, then, the findings of the present study suggest that absolute tuning perception is a replicable phenomenon but may not factor into situations in which relative pitch changes are present (i.e., virtually all musical settings outside of an experimental context).

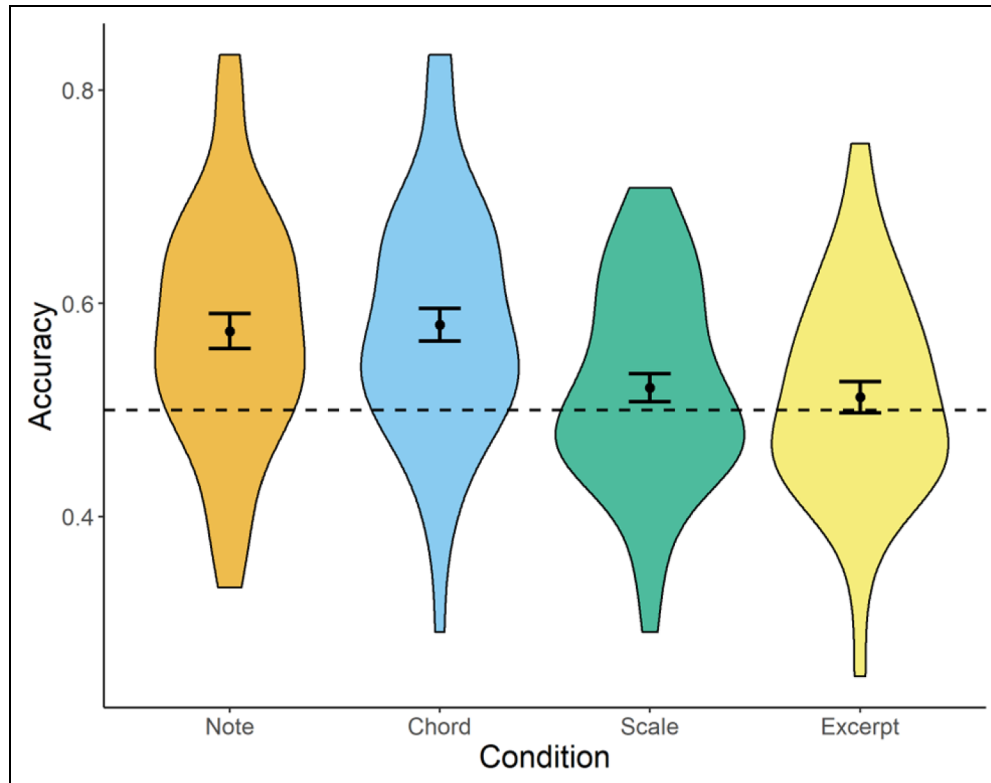


Figure 1. Violin plots of tuning judgment accuracy across all conditions.

Note: The dashed horizontal line represents chance performance. Error bars represent plus or minus one standard error of the mean. Note and Chord Conditions were robustly above chance ($p < .001$), whereas Scale and Excerpt Conditions did not significantly differ from the chance estimate.

Table 1. Correlation matrix from the experiment.

	1.	2.	3.	4.	5.	6.	7.	8.
1. Tuning Accuracy	—							
2. Age	.025	—						
3. Education	-.054	-.026	—					
4. Music Skill	.071	.193**	-.016	—				
5. Pitch Perception	.019	.134	.049	.393***	—			
6. Music Appreciated	-.064	.075	-.020	-.445***	-.263***	—		
7. Music Encouraged	-.017	.064	-.011	-.334***	.241***	.737***	—	
8. Music Training (y/n)	.086	.100	-.081	-.570***	.199**	.259***	.248***	—

Note: Values represent Pearson correlation coefficients. Significant values are bolded. ** $p < .01$ *** $p < .001$.

These findings can inform our understanding of pitch representation more broadly. Notably, the data reaffirm that pitch memory representations are context dependent (Hoeschele, 2017) and, furthermore, that absolute and relative pitch do not exist as entirely separate mechanisms of auditory processing. Everyday listeners (i.e., without AP and not specifically recruited for their musical expertise) seem to have pitch memory that is shaped by both absolute and relative pitch cues. Although listeners can use *latent* absolute pitch memory (cf. Levitin & Rogers, 2005) for single notes or chords, relative pitch rapidly replaces absolute pitch as the dominant mechanism once introduced. The relative weighting of these cues seems to vary across

circumstances as a function of ecological validity of the musical stimulus: the more salient the relative pitch cues are in the stimulus, the greater the expected reliance on relative pitch. While it is clear that most listeners can make absolute judgments that are conceptually aligned with those made by AP possessors, these judgments are critically limited to a subset of simple musical stimuli and are easily overpowered by the presence of even the most basic relative pitch cues.

The present findings can also potentially inform our understanding of the phenomenon of AP. Although absolute pitch representations are, by definition, grounded in an explicit recognition and labeling of pitch chroma (and

thus might be considered context *independent*), the specific way in which absolute and relative pitch interact among AP possessors remains debated. For example, Miyazaki (1993) found that AP possessors have difficulty naming intervals if both notes are detuned by the same amount, suggesting an impoverished sense of relative pitch. On the other hand, Dooley and Deutsch (2011) found that in a direct comparison between AP and non-AP possessors with musical training, AP possession was strongly associated with *increased* performance at interval identification tasks; furthermore, this performance was not attributed to cues that would benefit AP possessors specifically, such as knowledge of musical key. The findings of Dooley and Deutsch (2011) therefore suggest that AP possessors may have more developed relative pitch representations than those without AP (cf. Miyazaki, 1993).

Conceptually aligning with the notion that relative pitch processing is robust in AP possessors, Hedger et al. (2013) observed that absolute pitch categories can be slightly shifted when listening to an altered tuning system that maintains relative pitch relationships. Follow-up work demonstrated that listening to altered tuning can even lead to systematic misclassifications of notes (e.g., labeling a slightly sharp C as a C# after listening to music that had been detuned). While these studies, taken together, do not elucidate the precise ways in which absolute and relative pitch cues operate among AP possessors, the present research points to the possibility of individual and situational variability in the extent to which listeners rely on absolute versus relative pitch cues, which may explain the diversity of conclusions in prior work on the matter. Indeed, the current findings are quite consistent with these two earlier studies suggesting that salient relative pitch cues can disrupt absolute pitch processing (Hedger et al., 2013; Van Hedger et al., 2018a), suggesting that relative pitch can substantially shape absolute pitch representations in a context-dependent manner, both for individuals with and without AP.

The present study has some limitations that could be addressed in future work. First, it is worth noting that this was an online study, which does not allow the same kind of participant supervision as in-lab research. However, we deliberately crafted several study parameters to mitigate the effects of the lack of supervision (e.g., initial pre-screening through CloudResearch, incorporation of attention checks). Perhaps most importantly, prior work using a similar paradigm (Van Hedger et al., 2017) recruited both online (MTurk) and in-lab participants and did not find significant differences between the two groups. More generally, other recent work has affirmed the validity of results from online auditory studies, and even points to their benefits, such as, for example, the ability to recruit significantly larger, more diverse populations and receive faster results efficiently (e.g., Eerola et al., 2021; Zhao et al., 2022). Indeed, the participants in this study range from 21 to 70 years in age – a significantly richer sample than could be easily achieved from conducting this

experiment in person. As such, we expect that our findings would hold just as well if the experiment were repeated in an in-person setting.

A second limitation is that the musical stimuli were, overall, relatively artificial. The stimuli with arguably the greatest ecological validity (the musical excerpts) were all taken from piano compositions from the mid-to-late nineteenth-century, written by European composers, and as such were relatively homogenous in terms of style. Although participants were asked to evaluate stimuli that ranged in complexity from isolated notes to short musical excerpts, all stimuli used only a piano timbre and were generated from MIDI files. Future work would benefit from an examination of how increasing the ecological validity of the tested sounds (e.g., expanding beyond piano, using actual recordings rather than MIDI files) might influence the results.

The finding that listeners use both absolute and relative pitch cues based on listening context has potential implications for investigating individual differences in pitch processing in future work. Given that absolute and relative pitch are theoretically independent (e.g., Ziv & Radin, 2014), researchers could calculate the relative strength of an individual's relative pitch, normalized against their absolute pitch, to more directly assess whether these modes of processing influence absolute tuning perception. For example, a testable prediction that stems from this treatment of absolute and relative pitch processing is that an individual who has strong relative pitch processing might be more likely to have disrupted absolute tuning judgments when relative cues are present (e.g., for scales and musical excerpts). Another testable and non-intuitive hypothesis is that AP possessors might not actually have intact absolute tuning perception across all musical stimuli. More specifically, if an individual has strong absolute pitch but also strong relative pitch, they might be more flexible and use relative pitch processing in situations where this is salient (e.g., listening to musical scales or excerpts). Thus, those individuals who display strong absolute and relative pitch processing might show attenuated absolute tuning judgments depending on the context, despite possessing AP.

Conclusion

Most listeners demonstrate above-chance absolute pitch memory, but the connection between absolute and relative pitch use, as well as the stability of each representation across different listening contexts, are not completely understood. The present experiment assessed whether perception of absolute tuning (i.e., whether a musical sound or group of sounds adhered to A440 tuning) persisted across multiple musical sounds, varying in ecological validity and availability of relative pitch cues. We found that listeners could reliably differentiate in-tune from out-of-tune musical sounds, but only when they did not contain relative pitch changes. Overall, then, these results suggest that absolute features of music related to tuning are remembered by

most listeners and can manifest in the form of explicit judgments, but these memory representations are subject to interference by more salient relative pitch cues.

Action Editor

Isabel Martinez, Universidad Nacional de La Plata, Facultad de Bellas Artes.

Peer Review

Kathrin Schlemmer, Katholische Universität Eichstätt-Ingolstadt; Sarah Sauvé, University of Lincoln, School of Psychology.

Contributions

SVH: Conceptualization, methodology, programming, data analysis, writing – original draft preparation and editing.

NRB: Data analysis, writing – original draft preparation and editing.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval

This study was approved by the University of Western Ontario Research Ethics Board (Protocol #112574).

All materials and data associated with this manuscript can be accessed via Open Science Framework (<https://osf.io/zjcvd/>).

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part through an institutional grant from the Huron University College Faculty of Arts and Social Sciences Research Grant awarded to SVH.

ORCID iDs

Stephen C. Van Hedger  <https://orcid.org/0000-0002-2448-9088>

Noah R. Bongiovanni  <https://orcid.org/0000-0002-1615-5352>

Note

1. Although the present experiment was not formally pre-registered, we opted to describe the correlational measures as exploratory because we did not have *a priori* predictions as to how either demographic or musical variables would relate to absolute tuning performance.

References

- Bachem, A. (1955). Absolute pitch. *The Journal of the Acoustical Society of America*, 27(6), 1180–1185. <https://doi.org/10.1121/1.1908155>
- Ben-Haim, M. S., Eitan, Z., & Chajut, E. (2014). Pitch memory and exposure effects. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 24–32. <https://doi.org/10.1037/a0033583>
- Calamassi, D., & Pomponi, G. P. (2019). Music tuned to 440 Hz versus 432 Hz and the health effects: A double-blind cross-over pilot study. *Explore (New York, N.Y.)*, 15(4), 283–290. <https://doi.org/10.1016/j.explore.2019.04.001>
- Creel, S. C., & Tumlin, M. A. (2012). Online recognition of music is influenced by relative and absolute pitch information. *Cognitive Science*, 36(2), 224–260. <https://doi.org/10.1111/j.1551-6709.2011.01206.x>
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavioral Research*, 47(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- Deutsch, D. (2013). Absolute pitch. In *The psychology of music* (pp. 141–182). Elsevier. <https://doi.org/10.1016/B978-0-12-381460-9.00005-5>
- Dooley, K., & Deutsch, D. (2011). Absolute pitch correlates with high performance on interval naming tasks. *The Journal of the Acoustical Society of America*, 130(6), 4097–4104. <https://doi.org/10.1121/1.3652861>
- Eerola, T., Armitage, J., Lavan, N., & Knight, S. (2021). Online data collection in auditory perception and cognition research: Recruitment, testing, data quality and ethical considerations. *Auditory Perception and Cognition*, 4(3–4), 251–280. <https://doi.org/10.1080/25742442.2021.2007718>
- Eitan, Z., Ben-Haim, M. S., & Margulis, E. H. (2017). Implicit absolute pitch representation affects basic tonal perception. *Music Perception: An Interdisciplinary Journal*, 34(5), 569–584. <https://doi.org/10.1525/mp.2017.34.5.569>
- Farnsworth, P. R. (1968). *The social psychology of music*. The Iowa State University Press.
- Hedger, S. C., Heald, S. L. M., & Nusbaum, H. C. (2013). Absolute pitch may not be so absolute. *Psychological Science*, 24(8), 1496–1502. <https://doi.org/10.1177/0956797612473310>
- Hoeschele, M. (2017). Animal pitch perception: Melodies and harmonies. *Comparative Cognition & Behavior Reviews*, 12, 5–18. <https://doi.org/10.3819/CCBR.2017.120002>
- Jakubowski, K., Müllensiefen, D., & Stewart, L. (2017). A developmental study of latent absolute pitch memory. *Quarterly Journal of Experimental Psychology*, 70(3), 434–443. <https://doi.org/10.1080/17470218.2015.1131726>
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11(3), 346–374. [https://doi.org/10.1016/0010-0285\(79\)90016-1](https://doi.org/10.1016/0010-0285(79)90016-1)
- Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: Perception, coding, and controversies. *Trends in Cognitive Sciences*, 9(1), 26–33. <https://doi.org/10.1016/j.tics.2004.11.007>
- Litman, L., Robinson, J., & Abberbock, T. (2017). Turkprime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 49(2), 433–442. <https://doi.org/10.3758/s13428-016-0727-z>
- Loui, P. (2016). Absolute pitch. In S. Hallam, I. Cross, & M. Thaut (Eds.) *The Oxford handbook of music psychology* (pp. 81–94). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198722946.013.10>
- Marsman, M., & Wagenmakers, E. J. (2017). Bayesian Benefits with JASP. *European Journal of Developmental Psychology*, 14(5), 545–555. <https://doi.org/10.1080/17405629.2016.1259614>

- Miyazaki, K. (1993). Absolute pitch as an inability: Identification of musical intervals in a tonal context. *Music Perception: An Interdisciplinary Journal*, 11(1), 55–71. <https://doi.org/10.2307/40285599>
- Miyazaki, K. (1995). Perception of relative pitch with different references: Some absolute-pitch listeners can't tell musical interval names. *Perception & Psychophysics*, 57(7), 962–970. <https://doi.org/10.3758/BF03205455>
- Miyazaki, K., Rakowski, A., Makomaska, S., Jiang, C., Tsuzaki, M., Oxenham, A. J., Ellis, G., & Lipscomb, S. D. (2018). Absolute pitch and relative pitch in music students in the east and the west: Implications for aural-skills education. *Music Perception: An Interdisciplinary Journal*, 36(2), 135–155. <https://doi.org/10.1525/mp.2018.36.2.135>
- Nees, M. A. (2016). Have we forgotten auditory sensory memory? Retention intervals in studies of nonverbal auditory working memory. *Frontiers in Psychology*, 7(DEC), 1–6. <https://doi.org/10.3389/fpsyg.2016.01892>
- Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, 98(1), 1–11. <https://doi.org/10.1016/j.cognition.2004.09.008>
- 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), 1–7. <https://doi.org/10.1111/j.1467-7687.2005.00387.x>
- Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science*, 14(3), 262–266. <https://doi.org/10.1111/1467-9280.03432>
- Smith, N. A., & Schmuckler, M. A. (2008). Dial A440 for absolute pitch: Absolute pitch memory by non-absolute pitch possessors. *The Journal of the Acoustical Society of America*, 123(4), EL77–EL84. <https://doi.org/10.1121/1.2896106>
- Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, 113(2), 345–361. <https://doi.org/10.1037/0033-2909.113.2.345>
- Terhardt, E., & Seewann, M. (1983). Aural key identification and its relationship to absolute pitch. *Music Perception: An Interdisciplinary Journal*, 1(1), 63–83. <https://doi.org/10.2307/40285250>
- Van Hedger, S. C., Bongiovanni, N. R., Heald, S. L. M., & Nusbaum, H. C. (2023). Absolute pitch judgments of familiar melodies generalize across timbre and octave. *Memory & Cognition*. <https://doi.org/10.3758/s13421-023-01429-z>
- Van Hedger, S. C., Heald, S. L. M., Huang, A., Rutstein, B., & Nusbaum, H. C. (2017). Telling in-tune from out-of-tune: widespread evidence for implicit absolute intonation. *Psychonomic Bulletin and Review*, 24(2), 481–488. <https://doi.org/10.3758/s13423-016-1099-1>
- Van Hedger, S. C., Heald, S. L. M., Uddin, S., & Nusbaum, H. C. (2018a). A note by any other name: intonation context rapidly changes absolute note judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 44(8), 1268–1282. <https://doi.org/10.1037/xhp0000536>
- Van Hedger, S. C., Heald, S. L. M., & Nusbaum, H. C. (2018b). Long-term pitch memory for music recordings is related to auditory working memory precision. *The Quarterly Journal of Experimental Psychology*, 71(4), 879–891. <https://doi.org/10.1080/17470218.2017.1307427>
- Vuvan, D. T., Podolak, O. M., & Schmuckler, M. A. (2014). Memory for musical tones: The impact of tonality and the creation of false memories. *Frontiers in Psychology*, 5, 582. <https://doi.org/10.3389/fpsyg.2014.00582>
- Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, and Psychophysics*, 79(7), 2064–2072. <https://doi.org/10.3758/s13414-017-1361-2>
- Zhao, S., Brown, C. A., Holt, L. L., & Dick, F. (2022). Robust and efficient online auditory psychophysics. *Trends in Hearing*, 26, 233121652211187. <https://doi.org/10.1177/23312165221118792>
- Ziv, N., & Radin, S. (2014). Absolute and relative pitch: Global versus local processing of chords. *Advances in Cognitive Psychology*, 10(1), 15–25. <https://doi.org/10.2478/v10053-008-0152-7>