Just Can't Get You Out of My Head: Associations Between Involuntary Musical Imagery, Singing Accuracy, and Music Perception

David J. Vollweiler & Peter Q. Pfordresher
University at Buffalo, SUNY

Involuntary musical imagery (INMI, or "earworms") are short, repetitive sections of music that spontaneously occur. Prior research observed significant positive correlations between self-reported music experience and frequency of earworms, and the use of earworms to enhance semantic memory. Previous hypotheses suggest that singing accuracy uses voluntary imagery to help people imitate pitches, and that singing accuracy is positively related to an individual's music perception abilities. However, we know of no research that has assessed the associations between INMI and non-self-reported musical abilities (specifically looking at singing accuracy and music perception). Participants ($N = 152$) completed tasks assessing their singing accuracy (the Seattle Singing Accuracy Protocol; SSAP), music perception (the Profile of Music Skills; Mini-PROMS), and a self-report measure of involuntary musical imagery (the Involuntary Musical Imagery Survey; IMIS). Here we show that earworms are associated with both singing accuracy and music perception. These results suggest that musical ability is associated with the quality of auditory imagery experiences, though not necessarily the frequency with which images spontaneously emerge into one's consciousness.

Received: June 6, 2023, accepted September 25, 2024.

Key words: earworms, involuntary musical imagery (INMI), singing accuracy, music perception, musical ability

The phenomenon of involuntary musical imagery, commonly called "earworms" has been a topic of keen interest in music psychology research for over a decade. Most people have the experience of a musical tune that gets "stuck in the head" (Liikkanen, 2011) and remains highly resistant to any attempt at removal (Beaman & Williams, 2010). Yet, individuals can vary considerably in their vulnerability to these experiences and the vividness of their earworm experiences (Bailes, 2007). In the present paper we explore two objective measures of musical ability that may predict individual differences in INMI experiences: music performance ability (measured using accuracy of pitch matching in singing), and music perception ability. Our exploration of this relationship is grounded in previous empirical and theoretical research linking these musical abilities with the vividness and accuracy of voluntary musical imagery (e.g., Greenspon & Pfordresher, 2019; Pfordresher, Halpern, & Greenspon, 2015; Pfordresher & Halpern, 2013).

Involuntary Musical Imagery (INMI) episodes, which are more commonly called earworms, can be defined as short, repetitive sections of music that spontaneously enter one's consciousness without effort (Jakubowski & Williamson, 2014). Previous research has focused on possible causes of earworms. For instance, music that has a fast tempo and a melodic contour that is frequently heard is more likely to yield earworms than music with a slow tempo and a less common melodic contour (Jakubowski et al., 2016). In addition to the acoustic structure of music, INMI can be elicited by a variety of state-based variables, as identified by Williamson and colleagues (2012). Examples include the experience of a positive emotion or nostalgia (see also Janata et al., 2007), the experiences of different moods and emotions, being stressed or surprised, and being in a state of low attentional arousal (Williamson et al., 2012). Bailes (2007) showed that the immediate situation an individual is in (e.g., interacting with others or working at various times throughout the day) can also lead to an increase in the number of INMI episodes and individual experiences. There is also evidence suggesting that specific personality traits may cause one to be more susceptible to INMI. For instance, individuals scoring high on neuroticism report greater earworm frequency than other individuals (Floridou et al., 2012; Kellaris, 2003). Not surprisingly, physical exposure to a musical stimulus can also elicit subsequent involuntary imagery of that stimulus. INMI can also be elicited by association with the recollection of associated memories (Williamson et al., 2012).

Music Perception, pp. 1–12, issn 0730-7829, electronic issn 1533-8312. © 2025 by the regents of the university of california all rights reserved. please direct all requests for permission to photocopy or reproduce article content through the university of california press's Reprints and Permissions web page, https://online.ucpress.edu/journals/pages/reprintspermissions. DOI: https://doi.org/10.1525/mp.2025.2330968

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

Prior studies have explored the relationship between musical abilities and INMI experiences, but have not specifically focused on singing ability, resulting in mixed findings. Studies have failed to find associations between self-reported music training and INMI (Beaman & Williams, 2010; Müllensiefen et al., 2014), no association (Weir et al., 2015), or mixed effects (Floridou et al., 2015). The present research probes whether trait-based variables associated with musical ability are also associated with one's susceptibility to earworms, as well as the quality of experienced earworms. Specifically, we address whether the frequency and/or vividness of earworms correlates with accuracy of pitch matching during singing (a.k.a. vocal pitch imitation ability), and musical perception ability. To test this idea, we used measures of earworms as a novel test for a hypothesis concerning the mechanisms that underlie vocal pitch imitation and can account for deficits like poor-pitch singing.

Given that previous studies have shown an association between singing accuracy and voluntary imagery, we addressed whether a similar association exists between singing accuracy and INMI. Previous evidence is mixed concerning what outcome we may anticipate. On the one hand, both voluntary and involuntary imagery may draw on memory resources. Whereas Baddley and Andrade (2000) provide evidence that the vividness of voluntary imagery depends on activation in working memory, Kubit and Janata (2022, 2023) reported that involuntary imagery episodes help consolidate long-term memories. If both forms of imagery are integrated with memory, we may anticipate associations between them. Likewise, ratings of INMI frequency were found to correlate with self-assessment of the vividness of auditory imagery using the BAIS (Floridou et al., 2015). On the other hand, Loutari et al. (2023) found that individuals with congenital amusia (a deficit of perception that is often comorbid with poor-pitch singing) reported deficient imagery vividness on the BAIS yet did not differ from normal listeners with respect to INMI frequency. We anticipated finding an association between singing accuracy and INMI frequency based on the hypothesized role of imagery in sensorimotor associations that underlie singing (Pfordresher, Halpern, & Greenspon, 2015), and the apparent use of a common memory substrate.

Whereas previous studies of INMI have tended to focus on INMI frequency and its hedonic value, we also assessed measures that refer to the vividness of INMI; that is, the degree to which INMI experiences resemble veridical perceptual experiences. Although these items are part of the original Involuntary Musical Imagery Scale (IMIS), previous analyses of this measure have focused on other items (Floridou et al., 2015). Our incorporation of vividness here was based on the aforementioned associations of the vividness of voluntary imagery with singing accuracy (e.g., Pfordresher & Halpern, 2013).

In addition to singing accuracy, we also assessed music perception ability. Previous research suggests that singing accuracy is associated with music perception ability, though less strongly (or not at all) with simple pitch discrimination (Pfordresher & Brown, 2007; Pfordresher & Nolan, 2019). By assessing singing accuracy and music perception, we aimed to evaluate how both perception and production ability may jointly or independently predict INMI experiences.

We focus on the vividness and frequency aspects of INMI experiences in the present study. Prior literature has assessed vividness of voluntary auditory imagery (Greenspon & Pfordresher, 2019; Pfordresher & Halpern, 2013), as already noted, but we know of no studies that focus on vividness of INMI. Likewise, prior literature has found significant correlations between self-reported music training and frequency of INMI (Likkannen, 2011), but we do not know of any studies that used tasks that assessed musical abilities when looking at the associations with INMI. Therefore, we predict that the frequency and vividness of earworms would be associated with accurate performance on singing and music perception tasks.

Method

PARTICIPANTS

Participants (N = 226) were recruited from the Introduction to Psychology subject pool at the University at Buffalo, SUNY. Participants were 107 females and 119 males, aged 18–50 (M = 20.13, SD = 3.22). The average years of music experience for participants was 5.21 (SD = 5.57). The average time participants reported having played an instrument was 7.24 years (SD = 6.21), and the average amount of participants' vocal (singing) experience was 6.29 years (SD = 5.11). Seventy-six percent of participants (N = 172) reported speaking English as their primary language. The Appendix lists the break-down of native languages across all participants. All participants self-reported normal hearing.

Participants were removed during the analysis process if they did not know what an earworm was (N = 71). The final sample (N = 152) included 69 females and 83 males, aged 18–50 (M = 20.14, SD = 3.09). The average years of music experience for

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

participants was 5.75 years (*SD* = 5.78), with a range of 0 to 40 years. Seventy-nine percent of participants (*N* = 121) reported speaking English as their primary language. Compensation for completing this study included partial course credit.

APPARATUS

Participants were seated in an acoustically treated sound booth (Whisper Room Inc., whisperroom.com) where they were asked to complete vocal productions. Auditory stimuli were presented over Sennheiser HD 280 Pro headphones. Vocal performance was recorded using a Shure SM58 microphone connected through a Lexicon Omega I/O box.

MATERIALS

Involuntary Musical Imagery Scale

The Involuntary Musical Imagery Scale (IMIS; Floridou et al., 2015) is a self-report questionnaire that takes approximately 20 minutes to complete. It assesses information based on nine hypothesized dimensions of earworms including: frequency and intensity, potential for disturbance, trigger sensitivity, psycho-physiological match, vividness, induced emotions, approach or withdrawal, familiarity, and the time and length of the earworms. The *Frequency and Intensity* dimension assesses how often a person experiences earworms and when they occur throughout the day. The *Disturbance* dimension of earworms assesses how easy or difficult it is to get rid of an earworm and how irritating and distracting an earworm is. The *Trigger Sensitivity* dimension assesses the activity that causes the earworm (e.g., thinking of past or future events, exercising, engaging in mental activities, doing routine activities, etc.). The *Psycho-Physiological* dimension assesses whether an earworm matches one's mood, stress, or movement. The *Vividness* dimension assesses how easily an earworm can be brought into consciousness. The dimension of *Induced Emotions* assesses whether the earworm is pleasant, energizing, boring, or annoying. The *Approach or Withdrawal* dimension assesses an individual's reaction to an earworm. The *Familiarity* dimension assesses whether or not the earworm is related to music that the individual heard previously. The last dimension, *Time and Length* of earworms, assesses the duration of an earworm and how long it is stuck in one's head. Table 1 shows examples of INMI statements used in the IMIS.

All responses to items (except for those from the time and length dimension) used a five-point Likert scale in which a response of "1" indicates that participants always experienced a scenario, "3" indicates that participants sometimes experienced a scenario, and "5" indicates that participants never experienced a scenario.

TABLE 1. *IMIS Dimensions and Examples*

	Example of IMIS Statement
Frequency and Intensity	I experience earworms.
Potential for Disturbance	I try hard to get rid of my earworms
Trigger Sensitivity	I am unaware of what caused an earworm
Psycho-Physiological Match	My earworms don't necessarily match my mood
Vividness	My earworms are not as vivid as hearing real music
Induced Emotions	The experience of earworms is not pleasant
Approach or Withdrawal	My typical reaction to an earworm is to do nothing
Familiarity	My earworms contain music that I have never heard before
Time and Length	On average, my earworm (the section of music that is stuck) lasts

Note: Bolded items represent the overarching categories used in the analyses below.

We chose to use the IMIS (Floridou et al., 2015) instead of earlier INMI questionnaires (e.g., Beaman & Williams, 2010) because the questions asked covered a larger array of topics pertaining to earworms. Some of these additional topics covered in the IMIS but not included in the earlier questionnaires are: trigger sensitivity, vividness, and familiarity. We were primarily interested in the frequency of earworms and vividness questions.

Seattle Singing Accuracy Protocol

The Seattle Singing Accuracy Protocol (SSAP; Demorest et al., 2015; Pfordresher & Demorest, 2021) is a standardized measure of singing accuracy based on a battery of tasks that lasts approximately 20 minutes. Code and audio files for the SSAP are available at osf.io/aynmv/. The SSAP consists of measures designed to assess an individual's singing accuracy, pitch perception, and musical background. We will be focusing on the pitch imitation tasks and musical background questionnaire. In the pitch imitation tasks, participants are asked to listen to a target stimulus and reproduce it by singing. The first group of ten trials consist of imitating a single vocal pitch. The next group of ten trials consist of imitating a single piano pitch. The last group of six trials require participants to imitate four-note vocal pitch sequences. The vocal pitch stimuli for the first and third groups of trials are based on recordings of a male or

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

female trained singer, in which the vocal gender was matched to the participant. Participants also completed a musical background questionnaire. This questionnaire asks participants about their demographic information and background in musical activities (e.g., "What was your primary instrument?", "How many years for your primary area?", and "How many years were you a part of your primary music group?").

Profile of Music Skills

The Profile of Music Skills (Mini-PROMS; Law & Zentner, 2012; Zentner & Strauß, 2017) is a 15-minute task comprising four subtests: Melody, Tuning, Accent, and Tempo. In all four subtests, participants hear a probe stimulus twice, followed by a comparison stimulus once. Participants are then asked to identify if the comparison stimulus was the same as or different than the probe stimulus. The Melody subtest assesses the participant's ability to determine whether a probe stimulus that is either a tonal or an atonal melody was the same as or different than a comparison stimulus that may include one or more altered pitches. The Tuning subtest assesses whether participants can identify whether the notes in a probe chord were the same as a comparison chord or whether one pitch in the comparison chord was mistuned (either sharp or flat). The Accent subtest assesses skills in discerning whether the emphasis of certain notes of a rhythmic pattern in a probe stimulus matched the pattern of emphasis in a comparison stimulus. The Tempo subtest assesses whether participants could tell if the speed of a comparison melody changed by either getting faster or slower from a probe melody.

We chose to use the Mini-PROMS, instead of other music perception tasks such as the Montreal Battery of Evaluation of Amusia (Peretz et al., 2003; Vuvan et al., 2018) or Gordon's Musical Aptitude Profile tests (Gordon, 1965) because it covers similar tasks (e.g., identifying differences in rhythm, tempo, and tuning) in a shorter period (15 minutes compared to 90 minutes). The Mini-PROMS was also chosen for the higher test-retest reliability measures (r > .85; Zentner & Strauß, 2017) compared to the MBEA reliability (r = .75; Peretz et al., 2003).

PROCEDURE

Participants first completed the SSAP, including the warm-up exercises, the pitch matching tasks, pitch discrimination task, and demographic questionnaire. After the SSAP was completed, participants completed the Mini-PROMS. Finally, participants completed the IMIS. At the beginning of this survey, we added the definition of an earworm to ensure that participants understood what an earworm is, and an attention check that asked participants to explain, in their own words, what an earworm is to ensure that they read and understood the definition of an earworm. A fixed order was used as we were primarily interested in individual differences across participants, and also because we had *a priori* reasons to believe that the IMIS was more likely to generate a carryover effect such that participants may not perform to the best of their abilities on the SSAP and the Mini-PROMS assessments if they already had an earworm stuck in their head when compared to starting with one of the other tasks.

DATA ANALYSIS

The data and code for this paper are available on the project's Open Science Framework page: osf.io/t2s73/. Pitch extraction and analysis is carried out automatically in the SSAP (Pfordresher & Demorest, 2020). Pitch information is extracted by using the Yin algorithm (de Cheveigné & Kawahara, 2002) on MATLAB. To analyze the pitch imitation tasks, the difference between the target pitches and the imitated pitches is estimated on a note-by-note basis. To reduce the problem of "scoops" at the beginning or end of the imitated notes, the pitch analyses are limited to the central portion of pitches (the interquartile range of all extracted samples) and use the median f_0 rather than the mean to limit the influence of outliers. The mean difference between target and imitated f_0 values (in cents) yields the measure *Absolute Pitch Deviation* across trials. Another measure, *Pitch Error*, arises from categorizing every absolute deviation greater than 50 cents (half a semitone) as an error and the remainder as correct. To calculate the pitch error for each participant, we calculated the percentage of mean total incorrect notes and report it as a decimal.

Data analyses were performed via the following stages. First, we evaluated bivariate correlations across all variables of interest. At this stage we looked at separate items in the IMIS and focused specifically on the *vividness* and *frequency* subscales. The reason for this was that the original IMIS included subsections for Vividness but did not include a way to score the section overall. We used these associations as the basis for a later principal components analysis to reduce the information from these large number of correlations for subsequent multiple regression. We also focused on individual scales within the PROMS (and on aggregated scores across all imitation tasks from the SSAP). While the IMIS (Floridou et al., 2015) does have an already established factor structure, the four factors (Negative valence, Movement, Personal reflections, and Help) do not include items on either the vividness subscale

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

questions or the frequency subscales that were also developed in the original IMIS. Consolidation of the PROMS was carried out simply by adding the performance across the subscales together to create an overall score. Information-reduced measures were then used to probe interrelations across the three main domains of interest: INMI experience, singing accuracy, and music perception ability. We adopted multiple regression to assess these interrelationships.

Results

We predicted that the frequency and vividness of earworms would be positively related to singing accuracy and music perception. To look at these associations, we correlated scores from individual statements on the IMIS with a person's singing accuracy score or accuracy of music perception scores in each of the PROMS subtest (see Table 2 for Descriptive Statistics). Missing data were dealt with by omitting participants from individual correlation analyses; this resulted in the loss of no more than five participants from any one correlation.

TABLE 2. *Descriptive Statistics for Musical Abilities*

	Mean	*SD*	Range
Pitch Error	.34	.25	0 – 1
Absolute Pitch Deviation	79.38	108.93	11.04 – 1069.79
Melody Subtest of Mini-PROMS	5.44	1.68	1 – 9.50
Tuning Subtest of Mini-PROMS	4.47	1.45	.50 – 8
Accent Subtest of Mini-PROMS	5.33	1.72	0 – 9.50
Tempo Subtest of Mini-PROMS	5.41	1.46	1 – 8
Overall Score on Mini-PROMS	20.65	4.67	3 – 32.50

INMI AND SINGING ACCURACY

First, we looked at the correlations between singing accuracy and IMIS items assessing the *Vividness* dimension of earworms (where a score of 1 would represent always having this experience, and a score of 5 would represent never having this experience). All the correlations between singing accuracy and the vividness of earworms can be found in Table 3. Reverse coded items on the vividness subscale of the IMIS can also be found in Table 3. These items gauge how closely features of the earworm match acoustic properties of the source. For both pitch error and absolute pitch deviation, higher scores indicate that a participant performed worse at a given task. We found significant positive correlations between vividness of earworm tempo with pitch error (Figure 1A) and absolute pitch deviation (Figure 1B), indicating that poorer reported vividness (high IMIS values) corresponded in general to less accurate pitch imitation (higher errors/deviations). There were also significant positive correlations between pitch error and earworm timbre, via the statement, "if my earworm contains an instrument (e.g., a trumpet) then it sounds very much like the original" (Figure 1C). Another significant timbre-based association was found between pitch error and the IMIS statement, "If my earworm contains singing the voice/s sounds very much like the original version" (Figure 1D). As the stronger associations for singing accuracy arose from correlations with pitch errors, we used this measure for subsequent analyses involving singing.

Contrary to one of our predictions, we did not find any correlations between singing accuracy and any items in the *Frequency and Intensity* dimension of the IMIS. The *p* values associated with all correlations between measures of pitch accuracy (error and deviation scores) and items assessing the frequency of earworms were all $p > .20$.

TABLE 3. *Correlations with Singing Accuracy and Vividness of Earworms*

	Pitch Error	Abs. Pitch Deviation
My earworms are not as vivid as hearing real music[a]	$-.10^{1}$	$-.16^{3}$
The speed of my earworm matches the speed of the original music	$.23^{**2}$	$.24^{**4}$
If my earworm contains lyrics, then the singing voice I hear sounds like my own voice (rather than the original)	$.07^{1}$	$.08^{3}$
If my earworm contains many instruments, I'm not able to hear them all at the same time[a]	$-.08^{1}$	$-.16^{3}$
The lyrics in my earworms are not accurate[a]	$-.03^{2}$	$-.12^{4}$
If my earworm contains an instrument (e.g., a trumpet) then it sounds very much like the original	$.21^{**2}$	$.15^{3}$
If my earworm is a song, I experience only the tune without the words	$.03^{1}$	$.09^{3}$
If my earworm contains singing the voice/s sounds very much like the original version	$.19^{*2}$	$.06^{3}$

* $p < .05$, ** $p < .01$, *** $p < .001$; [a] = Item that has been Reverse Coded (Scale: 1 = Never, 5 = Always); [1] df = 149, [2] df = 148, [3] df = 144, [4] df = 143

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

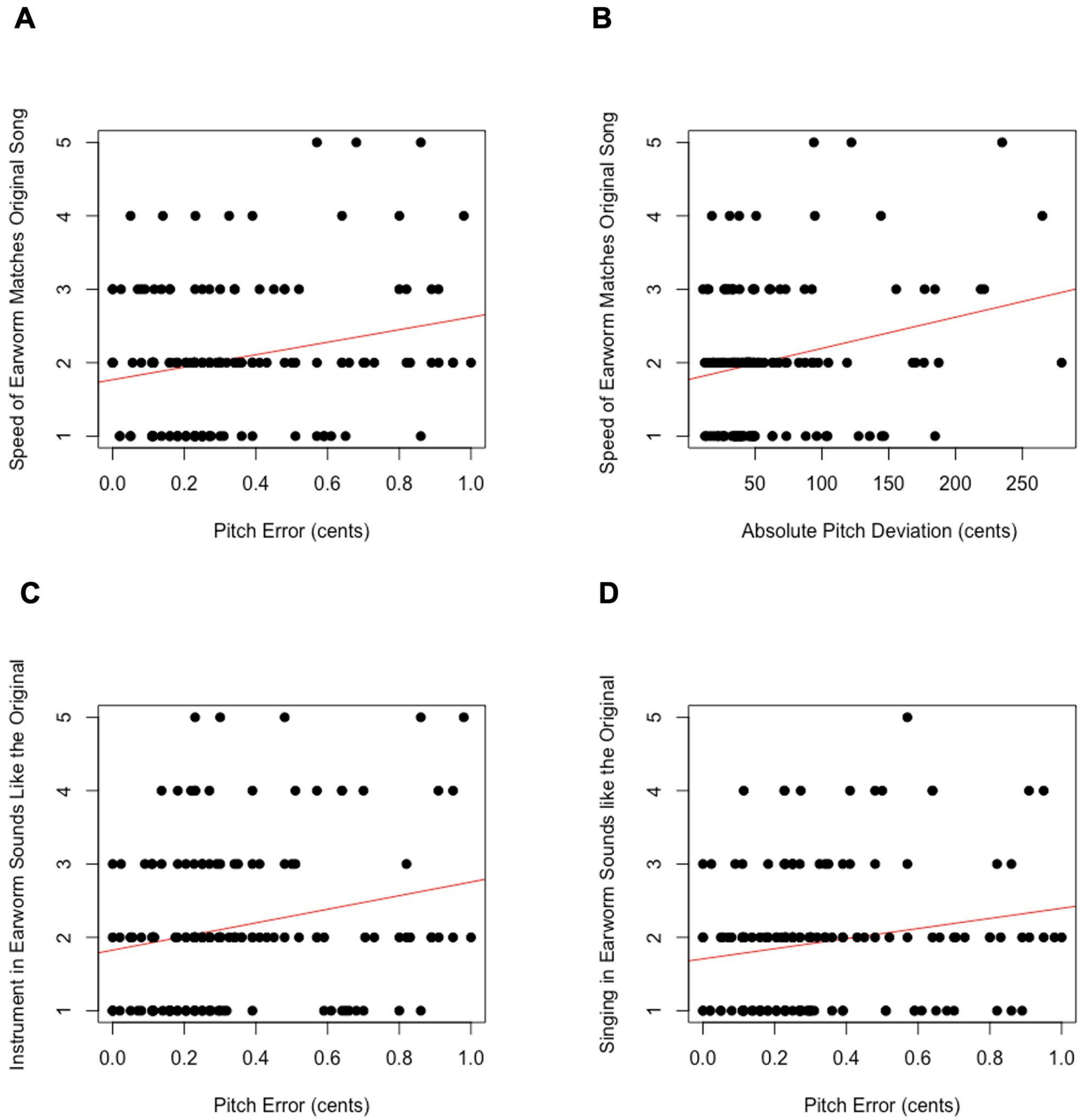

FIGURE 1. Singing accuracy and vividness of earworm scatterplots. Note: All of the associations in the above figures were predicted to be positive associations. The scale on the *y*-axis for all the figures is 1 = *Always experiences a scenario* and 5 = *Never experiences a scenario*.

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

INMI AND MUSIC PERCEPTION

Next, we looked at the correlations between accuracy of music perception (based on the scores from the Mini-PROMS) and IMIS items assessing the *Vividness* dimension of earworms (where a score of 1 would represent always having this experience, and a score of 5 would represent never having this experience). All the correlations between the accuracy of music perception and the vividness of earworms can be found in Table 4. Negatively worded items within the Vividness subscale of the IMIS have been reverse scored and are labeled in Table 4. For all items, we would expect a negative

TABLE 4. *Correlations between Accuracy of Music Perception and Vividness of Earworms*

	Overall	Melody	Tuning	Accent	Tempo
My earworms are not as vivid as hearing real music[a,1]	−.09	−.04	−.09	−.04	−.11
The speed of my earworm matches the speed of the original music[2]	−.17*	−.18*	−.04	−.12	−.15
If my earworm contains lyrics, then the singing voice I hear sounds like my own voice (rather than the original)[1]	.17*	.16	.06	.12	.18*
If my earworm contains many instruments, I'm not able to hear them all at the same time[a,1]	−.12	−.12	−.02	−.10	−.10
The lyrics in my earworms are not accurate[a,2]	−.13	−.07	−.12	−.09	−.10
If my earworm contains an instrument (e.g., a trumpet) then it sounds very much like the original[1]	−.22**	−.20*	−.15	−.17*	−.11
If my earworm is a song, I experience only the tune without the words[1]	.05	.03	.11	.02	.02
If my earworm contains singing the voice/s sounds very much like the original version[2]	−.14	−.07	−.03	−.14	−.17*

* $p < .05$, ** $p < .01$, *** $p < .001$; [a] = Item that has been Reverse Coded (1 = Never, 5 = Always); [1] df = 149, [2] df = 148

TABLE 5. *Correlations between Accuracy of Music Perception and Frequency of Earworms*

	Overall	Melody	Tuning	Accent	Tempo
I Experience Earworms	−.07	−.02	.03	−.04	−.20*
If I have an earworm, it appears once throughout the day	.08	−.02	.24**	< .01	.04
I get earworms soon after waking up	.04	.11	.01	−.04	.02
I get earworms at the end of the day	−.01	.05	.06	−.08	−.04
I get the same earworm coming back again and again	.04	−.01	.12	−.05	.05

* $p < .05$, ** $p < .01$, *** $p < .001$; df = 149

relationship between vividness of INMI and music perception ability given that high values on the PROMS indicate accurate performance. There were significant negative correlations between the accuracy of music perception and the vividness of an earworm with speed and whether the instrument in the earworm sounds like the original music. Because more INMI items were correlated with the overall perception score than with scores on individual subtests, we used the overall score in subsequent analysis comparing music perception to INMI and singing accuracy.

We next address the association between accuracy of music perception and items assessing the frequency of earworms (Table 5). Items on the frequency subscale of the IMIS were scored such that a score of "1" would represent always having this experience, and a score of "5" would represent never having this experience. This scoring was done post hoc. Similar to singing accuracy, most correlations were nonsignificant, with the exception of the negative association between the tempo sub-score of the Mini-PROMS and the statement, "I experience earworms," and a positive association between the tuning sub-score of the Mini-PROMS and the statement, "If I have an earworm, it appears once throughout the day." The positive association could be due to how we interpreted the IMIS statement as it is a two-tailed item (where the response of "never" could be that they experience at least two earworms per day or zero earworms per day). This suggests that people who are performing well on the tuning subscale of the PROMS are experiencing at least two earworms per day (since we see across all of the tasks that those who perform better on the task have more vivid and frequent episodes of INMI).

INMI AND MUSIC TRAINING

We also looked at the correlations between years of private music training and IMIS items assessing the *Vividness* dimension of earworms. We found no significant correlations between the vividness of INMI and years of music training. The p values associated with the correlations between music training and items assessing the vividness of earworms were all $p > .11$. Likewise, we found no significant correlation between years of music training and the *Frequency and Intensity* dimension of the IMIS. The p values associated with these correlations were all $p > .05$.

INFORMATION REDUCTION FOR INMI

Because both singing and perception yielded significant correlations with vividness, a principal components extraction with varimax rotation was then performed

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

on the eight vividness questions on the IMIS for a sample of 148 participants who successfully completed all the questions in the vividness subsection of the IMIS by using the "fa" command in the psych package in RStudio (version 2021.9.2.382; Revelle, 2023). This was used to estimate the number of factors, presence of outliers, absence of multicollinearity, and factorability of the correlation matrices. The goal of choosing the varimax rotation was to maximize the variance of the loadings within factors, across variables. The spread in the loadings is maximized—loadings that are high after extraction are higher after rotation and loadings that were low after extraction become lower after rotation. Varimax also tends to reapportion variance among the factors so that they become relatively equal in how much variance they account for in the observed variables.

Three components were extracted. As indicated by the squared multiple correlations (Table 6), all factors were internally consistent and well defined by the variables; the lowest of the SMCs for factors from variables was .40. Loadings of variables on factors, communalities, and percent of variance and covariance are shown in Table 7. Variables are ordered and grouped by size of loading to facilitate interpretation. Loadings under .26 (20% of variance) are replaced by zeros. The interpretative labels are Accuracy (Factor 1), Lyrics (Factor 2), and Acoustic Information (Factor 3).

In sum, the three factors of the vividness questions on the IMIS for this sample were Acoustic Match (e.g., if earworms contain either an instrument or singing, it sounds like the original version), Lyrics (e.g., If my earworm is a song, I experience only the tune without the words), and Accuracy (e.g., My earworms are not as vivid as hearing real music).

When comparing the IMIS statements that had significant correlations with singing accuracy, we found that the three vividness items yielding significant correlations with singing accuracy and perception all loaded onto the same factor (see bolded items in Table 7). Therefore, the analyses of relationships across the three domains of behavior under investigation (IMIS, singing accuracy, music perception), which we report next, use Factor 1 scores for a collective IMIS measure.

RELATING PERCEPTION, SINGING ACCURACY, AND MUSIC TRAINING AS PREDICTORS OF VIVIDNESS

Following information reduction, we addressed the relative contribution of singing accuracy and music perception in predicting the vividness of INMI experiences via multiple regression (Table 8) to examine the relative contribution of each variable. We also included music training as a predictor. Whereas both singing accuracy and music perception were significantly associated with INMI vividness in isolation, the regression model yielded a significant contribution of singing accuracy, but no significant contributions of music perception and years of music training. This result suggests that the ability to match pitch in production forms a stronger association with INMI than either music perception or music training.

TABLE 6. *SMCs for Factors with Variables as IVs with Orthogonal (Varimax) Rotation*

Squared Multiple Correlations		
Factor 1	Factor 2	Factor 3
.68	.47	.40

TABLE 7. *Principal Components Analysis for Vividness Statements*

	Factor 1[a]	Factor 2	Factor 3
If my earworm contains an instrument (e.g., a trumpet) then it sounds very much like the original			.59
If my earworm contains singing the voice/s sounds very much like the original version			.73
The speed of my earworm matches the speed of the original music			.50
My earworms are not as vivid as hearing real music[b]		.54	
The lyrics in my earworms are not accurate[b]		.52	
If my earworm contains many instruments, I'm not able to hear them all at the same time[b]	.37		
If my earworm is a song, I experience only the tune without the words	.34		
If my earworm contains lyrics then the singing voice, I hear sounds like my own voice (rather than the original)	.34		

Note: [a] Factor 1 = Accuracy; Factor 2 = Lyrics; Factor 3 = Acoustic Information; [b] Items have been reverse scored. The bolded items represent the items that yielded significant correlations with both pitch error rates and overall music perception scores.

Discussion

This paper focused on whether measures of musical performance (singing) and perception ability predict the frequency and vividness with which individuals experience involuntary musical imagery (INMI), often

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

TABLE 8. *Vividness of Earworms Regressed on Pitch Error, Music Perception and Years of Music Training*

IMIS Factor Scores for Acoustic Match		
	b [95% CI][a]	sr
Pitch Error	.71[.02, 1.41]*	.17
Music Perception	.03[-.01, .07]	.13
Years of Music Training	.01[-.02, .04]	.05

Note. sr = semi-partial correlations; * $p < .05$; ** $p < .01$; *** $p < .001$. [a] $R^2 = .02$, $F(3, 143) = 1.98$, $p = .12$

referred to as earworms. Our goal in doing this was to better understand the trait-based variables that may predispose individuals to more frequent and/or more vivid INMI experiences. We found that INMI vividness (specifically pertaining to the acoustic match), but not frequency, was positively associated with both singing accuracy and music perception, whereas no significant associations were found with self-reported INMI frequency. Whereas previous studies yielded mixed results on the relationship between INMI and musical ability when focused on self-reports of musical engagement, participation, and music training (Beaman & Williams, 2010; Beaty et al., 2013; Hyman et al., 2013; Liikkanen, 2011; Liptak et al., 2022; Müllensiefen et al., 2014; Williamson & Müllensiefen, 2012), the present data suggest that ability as measured in an active production task yields a more reliable association with vividness of INMI as it is an objective measure of musical abilities as compared to the self-reported amount of music training (see also Evans et al., 2024), which is more indirect.

The current focus on singing accuracy was inspired in part by research and theory suggesting that the vividness and accuracy of auditory imagery plays a critical role in singing accuracy (Greenspon & Pfordresher, 2019; Pfordresher & Halpern, 2013; Pfordresher, Halpern, & Greenspon, 2015; Silas et al., 2023). Importantly, previous studies focused on imagery as a predictor of singing. Here the focus was instead on imagery that emerges spontaneously and is typically not put into action via singing. As such, our research adopts regression designs to test the reverse relationship (singing ability as a predictor of imagery). In this context, we suggest that the present association arises as a byproduct of the sensorimotor network on which singing accuracy may be based (e.g., Berkowska & Dalla Bella, 2009; Hutchins & Moreno, 2013; Pfordresher, Demorest, et al., 2015; Welch, 1985; Zak et al., 2022). In other words, singing accuracy does not predict INMI because of overt singing activity during INMI, but rather the association is an indicator of a system in which those connections are well formed. It is important to note that because this study is correlational in nature, predictive relationships documented here do not necessarily imply causal relationships.

The association between INMI and singing accuracy found here mirrors previously reported associations between singing accuracy and voluntary imagery (Greenspon & Pfordresher, 2019; Pfordresher & Halpern, 2013; Silas et al., 2023), thus suggesting a similar basis for both forms of imagery. Some evidence suggests that both voluntary and involuntary musical imagery may draw on memory resources. For instance, Baddeley and Andrade (2000) provide evidence that the vividness of voluntary imagery depends on activation in working memory. Similarly, Kubit and Janata (2022, 2023) demonstrate that episodes of involuntary imagery help consolidate long-term memories. However, other studies have suggested that these forms of imagery may dissociate. For instance, congenital amusics report similar rates of INMI to controls while reporting less vivid voluntary imagery (Loutari et al., 2023), and self-reports of INMI frequency were not associated with individual differences in the accuracy of voluntary imagery among normal listeners (Weir et al., 2015). There are two potential reasons for these discrepancies. First, these earlier studies focused specifically on INMI frequency, which did not yield significant associations, and not on INMI vividness. Second, the use of singing accuracy as a common correlate for both types of imagery may have provided a source of individual variation that is more informative than those used in earlier studies. If there were individual differences in perception, we would have expected more robust correlations between vividness of INMI and music perception, and also a significant contribution to the multiple linear regression.

The present results show a dissociation between the frequency and vividness of INMI. Whereas significant to modest relationships were found between INMI vividness and measures of musical ability, no significant associations were found with INMI frequency. This was initially surprising to us, although on reflection these results mirror differences between frequency and vividness found when comparing INMI experiences across congenital amusics and controls (Loutari et al., 2023). This dissociation suggests that the fluid functioning of the sensorimotor network that forms associations between vocal imitation and imagery may influence the quality of spontaneously emitted auditory imagery, but not the likelihood of these images being generated. Based on previous research, it seems plausible that the generation of INMI may be based largely on environmental factors (such as previous exposure to an INMI-inducing tune) or individual state variables (such as focus of

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

attention, Williamson et al., 2012), whereas sensorimotor associations may not modulate the probability of the initial generation of INMI. Once generated, however, the sensorimotor network (a trait variable) may influence the veridicality of INMI in comparison to actual perception. This is a novel possibility from the current data and worth further exploration.

Some study limitations merit comment. One limitation of this study, and many others concerning INMI, is that it may be hard for participants to derive an accurate statistical sense of their experiences with earworms. Likewise, by obtaining self-report scores from the IMIS, the validity of the measure can vary depending on factors such as the honesty and accuracy of the respondent, the clarity of the questions, and the potential for bias or social desirability effects. In addition, some of the items on the vividness subscale of the IMIS (e.g., "'If my earworm contains lyrics, then the singing voice I hear sounds like my own voice (rather than the original)" and "If my earworm is a song, I experience only the tune without thewords") introduced some ambiguity when interpreting the results as to whether these items create veridical images and should have also been reverse coded for the analyses. These items were not reverse coded to maintain our *a priori* interpretation of the items that these items asked participants about veridical images they may have experienced. Future research designs may benefit from using experience sampling to obtain data over the course of time to get a better understanding of how frequently an individual experiences an earworm (Evans et al., 2024). Future research would also be able to determine how vivid an earworm is for an individual by asking how vivid the acoustic properties of an earworm are. Future research should look at how the association between voluntary musical imagery and singing accuracy may differ from the association between involuntary musical imagery and singing accuracy.

Another limitation of the present study is that the correlations with INMI, though statistically significant, were modest in size and account for a relatively small portion of variance. In addition, the removal of a significant portion of our sample due to being non-native English speakers, the possibility that the participants may have misread the definition, or individuals who don't experience INMI may have influenced robustness. These effect sizes are comparable to other associations with singing accuracy found in recent studies (e.g., Greenspon & Pfordresher, 2019; Pfordresher & Demorest, 2021; Pfordresher & Nolan, 2019), all of which likely relate to the complexity and multifaceted nature of musical abilities. The present research thus advances our understanding of INMI and musical ability but does not provide a complete picture. Future research should use experimental and/or longitudinal designs that may be able to tease apart the directionality of this relationship.

Conclusion

We sought to determine if there are any associations between the vividness of involuntary musical imagery and singing accuracy, and between the vividness of involuntary musical imagery and music perception. We used a self-report measure to determine frequency and vividness of earworms, and tasks designed to assess singing accuracy and music perception. We can conclude that the timbral vividness of an earworm is associated with both singing accuracy and music perception. Our findings offer confirmation that musical ability is associated with the quality of auditory imagery experiences, though not necessarily the frequency with which images spontaneously emerge into one's consciousness.

Author Note

David J. Vollweiler is now at the University of Nevada Las Vegas, Las Vegas, NV.

We have no known conflict of interest to disclose.

We thank Julie Aronsky, Thamarah Bouaz, Lihui De, Elizabeth Goldstein, and Richard Nicosia for assistance with data processing and data collection.

Correspondence concerning this article should be addressed to David J. Vollweiler (vollweil@unlv.nevada.edu).

References

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145.

Bailes, F. (2007). The prevalence and nature of imagined music in the everyday lives of musical students. *Psychology of Music*, *35*(4), 1–16.

Beaman, C. P., & Williams, T. I. (2010). Earworms (stuck song syndrome): Towards a natural history of intrusive thoughts. *British Journal of Psychology*, *101*(4), 637–653.

Beaty, R. E., Burgin, C. J., Nusbaum, E. C., Kwapil, T. R., Hodges, D. A., & Silvia, P. J. (2013). Music to the inner ears: Exploring individual differences in musical imagery. *Consciousness and Cognition*, *22*(4), 1163–1173.

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

Berkowska, M., & Dalla Bella, S. (2009). Acquired and congenital disorders of sung performance: A review. *Advances in Cognitive Psychology*, *5*(5), 69–83.

de Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, *111*(4), 1917–1931.

Demorest, S. M., Pfordresher, P. Q., Dalla Bella, S., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. F. (2015). Methodological perspectives on singing accuracy: An introduction to the special issue on singing accuracy (Part 2). *Music Perception*, *32*(3), 266–271.

Evans, M. G., Gaeta, P., & Davidenko, N. (2024). Absolute pitch in involuntary musical imagery. *Attention, Perception, and Psychophysics*, *86*, 2124-2135.

Floridou, G. A., Williamson, V. J., & Müllensiefen, D. (2012). Contracting earworms: The roles of personality and musicality. In E. Cambouropoulos, C, Tsougras, K. Mavromatis, K. Pastiadis (Eds.), *Proceedings of ICMPC-ESCOM 12*.

Floridou, G. A., Williamson, V. J., Stewart, L., Müllensiefen, D. (2015). The involuntary musical imagery scale (IMIS). *Psychomusicology: Music, Mind, and Brain*, *25*(1), 28–36.

Gordon, E. (1965). The Musical Aptitude Profile: A new and unique musical aptitude test battery. *Bulletin of the Council for Research in Music Education*, *6*, 12–16.

Greenspon, E. B., & Pfordresher, P. Q. (2019). Pitch specific contributions of auditory imagery and auditory memory in vocal pitch imitation. *Attention, Perception and Psychophysics*, *81*(7), 2473-2481.

Hutchins, S. & Moreno, S. (2013). The linked dual representation model of vocal perception and production. *Frontiers in Psychology*, *4*, 825.

Hyman, I. E., Burland, N. K., Duskin, H. M., Cook, M. C., Roy, C. M., & McGrath, J. C., & Roundhill, R. F. (2013). Going gaga: Investigating, creating, and manipulating the song stuck in my head. *Applied Cognitive Psychology*, *27*(2), 204–215.

Jakubowski, K., Finkel, S., Stewart, L., & Müllensiefen, D. (2016). Dissecting an earworm: Melodic features and song popularity predict involuntary musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(2), 122–135.

Jakubowski, K., & Williamson, V. J. (2014). Earworms. In W. F. Thompson (Ed.), *Music in the social and behavioral sciences: An encyclopedia* (*Vol. 1*, pp. 353–355). SAGE Publications Ltd.

Janata, P., Tomic, S. T., & Rakowski, S. K. (2007). Characterisation of music-evoked autobiographical memories. *Memory*, *15*(8), 845–860.

Kellaris, J. J. (2003). Dissecting earworms: Further evidence on the 'song-stuck-in-your-head' phenomenon. *Proceedings of the Society for Consumer Psychology Winter 2003 Conference* (pp. 220–222). American Psychological Society.

Kubit, B. M., & Janata, P. (2022). Spontaneous mental replay of music improves memory for incidentally associated event knowledge. *Journal of Experimental Psychology: General*, *151*(1), 1–24. https://doi.org/10.1037/xge0001050

Kubit, B. M., & Janata, P. (2023). Spontaneous mental replay of music improves memory for musical sequence knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(7), 1068–1090. https://doi.org/10.1037/xlm0001203

Law, L. & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PLOS One*, *7*(12), e52508.

Liikkanen, L. A. (2011). Musical activities predispose to involuntary musical imagery. *Psychology of Music*, *40*(2), 236–256.

Liptak, T. A., Omigie, D., & Floridou, G. A. (2022). The idiosyncrasy of involuntary musical imagery repetition (IMIR) experiences. *Music Perception*, *39*(3), 320–338. https://doi.org/10.1525/mp.2022.39.3.320

Loutrari, A., Ansell, K., Philip Beaman, C., Jiang, C., & Liu, F. (2023). Auditory imagery in congenital amusia. *Musicae Scientiae*, 102986492211228. https://doi.org/10.1177/10298649221122870

Müllensiefen, D., Fry, J., Jones, R., Jilka, S., Stewart, L., & Williamson, V. J. (2014). Individual differences predict patterns in spontaneous involuntary musical imagery. *Music Perception*, *31*(4), 323–338.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*(1), 58–75.

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "tone deafness." *Music Perception*, *25*(2), 95–115. http://doi.org/10.1525/MP.2007.25.2.95

Pfordresher, P. Q., & Demorest, S. M. (2020). Construction and validation of the Seattle Singing Accuracy Protocol (SSAP): An automated online measure of singing accuracy. In F. Russo, B. Ilari, & A. Cohen (Eds). *Routledge companion to interdisciplinary studies in singing* (Vol. 1, pp. 322–333). Routledge.

Pfordresher, P. Q., & Demorest, S. M. (2021). The prevalence and correlates of accurate singing. *Journal of Research in Music Education*, *69*(1), 5–23. https://doi.org/10.1177/0022429420951630

Pfordresher, P. Q., Demorest, S. M., Dalla Bella, S., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. F. (2015). Theoretical perspectives on singing accuracy: An introduction to the special issue on singing accuracy (part 1). *Music Perception*, *32*(3), 227–231. https://doi.org/10.1525/MP.2015.32.3.227

Pfordresher, P. Q., & Halpern, A. R. (2013). Auditory imagery and the poor-pitch singer. *Psychonomic Bulletin and Review*, *20*(4), 747–753. https://doi.org/10.3758/s13423-013-0401-8

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

Pfordresher, P. Q., Halpern, A. R., & Greenspon, E. B. (2015). A mechanism for sensorimotor translation in singing: The multi-modal imagery association (MMIA) model. *Music Perception*, *32*(3), 242–253.

Pfordresher, P. Q., & Nolan, N. P. (2019). Testing convergence between singing and music perception accuracy using two standardized measures. *Auditory Perception and Cognition*, *2*(1–2), 67–81. https://doi.org/10.1080/25742442.2019.1663716

Revelle, W. (2023). *psych: Procedures for Psychological, Psychometric, and Personality Research*. (2.3.3) [Computer Software]. https://cran.r-project.org/web/packages/psych/index.html

Silas, S., Müllensiefen, D., & Kopiez, R. (2024). Singing Ability Assessment: Development and validation of a singing test based on item response theory and a general open-source software environment for singing data. *Behavioral Research Methods*, *56*, 4358–4384. https://doi.org/10.3758/s13428-023-02188-0

Vuvan, D. T., Paquette, S., Goulet, G. M., Royal, I., Felezeu, M., & Peretz, I. (2018). The Montreal Protocol for Identification of Amusia. *Behavior Research Methods*, *50*(2), 662–672.

Weir, G., Williamson, V. J., & Müllensiefen, D. (2015). Increased involuntary musical mental activity is not associated with more accurate voluntary musical imagery. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 48–57.

Welch, G. F. (1985). A schema theory of how children learn to sing in tune. *Psychology of Music*, *13*(1), 3–18.

Williamson, V. J., Jilka, S. R., Fry, J., Finkel, S., Müllensiefen, D., & Stewart, L. (2012). How do "earworms" start? Classifying the everyday circumstances of Involuntary Musical Imagery. *Psychology of Music*, *40*(3), 259–284.

Williamson, V. J., & Müllensiefen, D. (2012) *Earworms from three angles: Situational antecedents, personality predisposition and the quest for a musical formula*. Paper presented at the ICMPC-ESCOM 12, Thessaloniki, Greece (pp. 1124–1132).

Zentner, M., & Strauß, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the Short-Proms and the Mini-Proms. *Annals of the New York Academy of Sciences*, *1400*(1), 33–45.

Zuk, J. M., Loui, P., & Guenther, F. H. (2022). Neural control of speaking and singing: The DIVA model of singing. *PsyArXiv*.

Downloaded from http://online.ucpress.edu/mp/article-pdf/doi/10.1525/mp.2025.2330968/874471/mp.2025.2330968.pdf by State University of New York at Buffalo, Peter Q Pfordresher on 15 May 2025

Appendix

NATIVE LANGUAGES OF PARTICIPANTS

	Number of Participants (% of participants)
English	172 (76.11%)
Amharic	1 (0.44%)
Urdu	1 (0.44%)
Swahili	1 (0.44%)
Chinese/Mandarin	18 (7.96%)
Cantonese	1 (0.44%)
Russian	2 (0.88%)
Ukranian	2 (0.88%)
Arabic	3 (1.33%)
Bangala	2 (0.88%)
Hindi	1 (0.44%)
Italian	1 (0.44%)
Nepali	1 (0.44%)
Telugu	1 (0.44%)
Wolof	1 (0.44%)
Korean	5 (2.21%)
Vietnamese	2 (0.88%)
Kurdish	1 (0.44%)
Spanish	2 (0.88%)
Amharic	1 (0.44%)
Bengali	5 (2.21%)
Greek	2 (0.88%)
Malayalam	1 (0.44%)