Is musical ability related to second-language acquisition? A meta-analysis.

Rachel M Thompson, Lauren K Salig, and L. Robert Slevc

Program in Neuroscience and Cognitive Science, University of Maryland, College Park

Author Note

1

Abstract

In our multicultural and interconnected world, the ability to learn new languages is important. However, there are significant differences in how successfully adults can learn aspects of non-native languages. Given robust relationships between musical ability and native-language processing, musical ability might also contribute to successful second-language acquisition. However, while several studies have assessed this relationship in various ways, the consistency and robustness of the relationship between musical ability and second-language learning remains unclear. Thus, we synthesized 191 effects across 60 independent studies (N = 3,462) with a robust variance estimation multivariate meta-analysis, and we narratively summarized partial correlation effects across 12 studies. The available evidence suggests that musical ability is indeed positively related to second-language learning, even after factoring in publication bias revealed by the meta-analysis. Although future work with more diverse participant populations and methodologies is needed to further disentangle this relationship, it is apparent that individuals with better musical ability are generally more successful at second-language learning.

Introduction

Humans have an impressive ability to learn the sounds and structure of their native linguistic and musical systems. However, this impressive ability to learn is diminished later in life, at least for language: Adults' ability to achieve native-like proficiency in a second language is notoriously variable, especially in the domains of phonology and morphosyntax (Flege et al., 1999; Golestani & Zatorre, 2009; Snow & Hoefnagel-Höhle, 1978).[1] While reaching native-like proficiency in a second language is not the end goal for all language learners, many may strive for native-like proficiency to avoid the social stigma associated with accents (Fuertes et al., 2012; Ramjattan, 2023) and to improve communication with native speakers (Christiner & Reiterer, 2015; Weber et al., 2011). Of course, this goal is met with varying success. Thus, two questions emerge: Why are some individuals more successful than others in learning a second language, and what characterizes successful second-language learners?

One idea, supported by a growing body of research, is that successful second-language acquisition is related to musical ability. This follows from broader links between musical and linguistic processing in the native language (for reviews, see e.g., Patel, 2008; Slevc, 2012), including evidence that individual differences in grammatical ability and speech processing are related to musical ability (e.g., Gordon et al., 2015; Jansen et al., 2023; Lee et al., 2020; Nayak et al., 2022). For example, musical ability predicts the distinctiveness of neural responses to closely related speech sounds (Parbery-Clark et al., 2012), speech discrimination performance (Swaminathan & Schellenberg, 2020), and sensitivity to prosody (Jansen et al., 2023; Morrill et al., 2015). One framework to understand these types of music-language relationships is Patel's (2014) OPERA hypothesis, according to which musical training promotes structural and functional changes in important auditory processing networks (cf. Kraus & Chandrasekaran, 2010). Alternatively, these relationships might not reflect effects of musical training, but rather exist because music and language both rely on similar underlying abilities (see, e.g., Mosing et al.,

---

[1] In contrast, there is some evidence that adults can successfully learn new musical systems (e.g., Loui, Wessel, & Hudson Kam, 2010; Pelofi & Farbood, 2021), but there has been relatively little work on this topic.

2014; Schellenberg, 2015). In either case, findings of associations between musical and native-language abilities are often taken as indication for deep, underlying relationships between music and language.

Although the specific reasons for music and language relationships are debated, it is relatively uncontroversial that musical and linguistic abilities are correlated. Indeed, recent meta-analyses have found evidence for relatively robust relationships between musical ability and native-language speech processing (Neves et al., 2022) and reading ability (Gordon et al., 2015). It is less clear, however, if this relationship extends to *second-language* learning. Some evidence suggests musical ability and musical training predict aspects of second-language proficiency, including perception and production of second-language phonology (Slevc & Miyake, 2006), perception of second-language prosody (see Jansen et al., 2023, for a meta-analysis), lexical tone discrimination (Cooper et al., 2017; Cooper & Wang, 2012), and speech segmentation (Gottfried et al., 2004; Lee & Hung, 2008). However, other evidence suggests that musical ability or training is unrelated to second-language learning (e.g., Boll-Avetisyan et al. 2016; Ning et al., 2015). For example, some studies have found no clear link between musical ability and vowel learning (Mokari & Werner, 2018), phonemic discrimination (Milovanov et al., 2010), and self-rated second-language abilities (Schellenberg et al., 2023). These varied results suggest a need for a formal meta-analytic assessment of the relationship between second-language abilities and musical experience/ability.

The Present Study

In light of these mixed results, we aimed to systematically assess existing findings on music and second-language relationships by conducting a meta-analysis on the relationship that musical ability (which here refers to performance on assessments of the perception or production of musical notes, melodies, or rhythms) and musical training (which here refers to the duration of prior musical experience) have with second-language learning. We also investigated several factors that might moderate such a relationship, including the type of musical measure used (e.g., ability assessment or self-reported training), the type of second-language learning assessed (e.g., lexical tones, syntax, or segmental

4

phonology), and learner characteristics (e.g., age or type of native language). As noted above, there are theoretical reasons to expect a link between music and second-language learning (e.g., Patel, 2014), and two narrative reviews of the relevant literature have concluded that there is, indeed, such a link (albeit a decade ago; Chobert & Besson, 2013; Zeromskaite 2014). However, there has not yet been a systematic meta-analysis of music and second-language relationships. Such a meta-analysis offers several benefits over narrative reviews, such as allowing for the quantification of effect sizes, assessment of potential moderating effects, and assessment of the quality of the literature (i.e., in terms of publication bias). This approach thus can offer a better understanding of the depth, breadth, and nuances of music–second-language relationships that we hope will help illuminate future research avenues.

Method

Search Strategy

The current meta-analysis was carried out in compliance with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA; Liberati et al., 2009). We performed a literature search across six databases, searched reference lists within relevant articles for studies, and consulted five review articles to select articles to include in analyses. Our search syntax for databases was: (second language OR foreign language OR bilingual) AND (musical ability OR musical experience OR musical aptitude OR musical training OR musicality), entered as a Boolean phrase. In total, we found 258 potentially relevant articles, of which 60 met our inclusion criteria. These 60 studies included 191 effect sizes and a total of 3,462 participants (see Figure 1 for the literature review process). The data and analysis scripts are available on the Open Science Framework (OSF): https://osf.io/4yzpq/.

Figure 1. *Flow chart of literature search and study selection.*
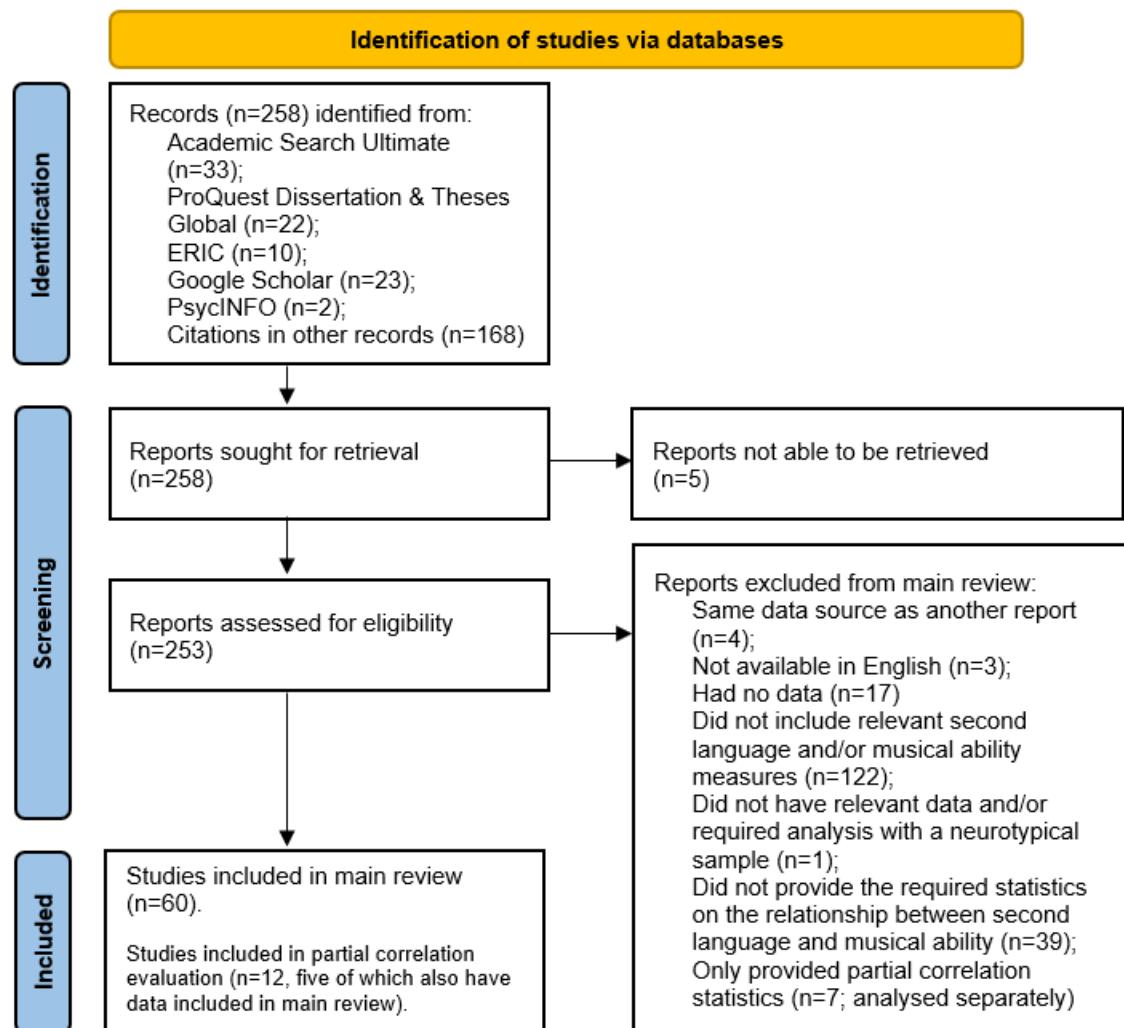
Note. The PRISMA diagram outlines the various steps of the literature search for study inclusion and exclusion. Reasons for exclusion are indicated at each step.

Inclusion Criteria

To be in the final analyses, studies must have met the following inclusion criteria:

1. Must be a published article or unpublished thesis/dissertation in English.

2. Must include a clear behavioral measure of musical aptitude, experience, or training. This could be a continuous measure based on scores from a test (e.g., Advanced Measures of Musical

Audition; Gordon, 1989), participants' self-report of musical training/experience, or groups defined from scores on a musical test.[2] Alternatively, experimenters could have provided musical training to participants as a form of a musical intervention.

3. Must include a clear behavioral measure of second-language learning/ability. This could include measures of second-language perception (e.g., discriminating between phonemes or lexical tones) and/or second-language production (e.g., repeating phonemes).

4. Must assess participants without identified neurological problems, learning impairments, or language impairments. If the sample included non-neurotypical participants, the study must also provide a relevant analysis including only neurotypical participants.

5. Must quantify the relationship between musical and second-language measures or predict second-language learning from musical measures via a Pearson correlation coefficient (r) or another statistic that could be transformed into a correlation coefficient (e.g., t, Cohen's d, or F; see *Converting Among Effect Sizes* below). Studies that reported partial correlations (such as findings from a multiple regression or partial correlation analysis) were excluded from the main analyses, but several of such studies are summarized narratively below (see *Narrative Summary of Partial Correlations*).

Converting Among Effect Sizes

Although most reported effect sizes were correlations, it was necessary to convert and transform effects that were reported as standard mean differences or with other effect size metrics. We used two web-based calculators to calculate and convert effect sizes: Campbell collaboration's online calculator (Wilson, n.d) to calculate effect sizes for standard mean differences and escal (Lin, n.d) to convert effect sizes to Pearson's r. We then converted all values to z scores via the Fisher *z*-transformation.

---

[2] We defined musical ability as some aspect of skill in perception or production of musical notes, melodies, tones, or rhythms, as measured by a test. We defined musical training as a measure of musical experience as a self-report measure. This could be treated as a continuous or a dichotomous measure.

Moderator Variables

A common challenge with meta-analyses is that studies on the same topic vary considerably. One way to address this issue is to consider specific dimensions on which studies systematically differ as potential moderators of an effect. As such, we identified (based on past work or on plausibility) six factors that seemed likely to relate to the strength of the observed music and second-language relationship.

1. Age (Children or Adults): Learning the novel sounds and structures of a second language seems to become more difficult as people age (Flege et al., 1999; Johnson & Newport, 1989), suggesting that musical ability – second-language relationships might also differ by age. In the studies examined here, participants' age varied from 5 to 69 years; however, the distribution was relatively bimodal: Studies generally included either relatively young (5-9 years old) or adult (18-69 years old) participants. We thus classified participant groups that were under 15 years old as *children* and groups over 16 years old as *adults*.

2. Musical Measure (Ability, Training, or Intervention): There is considerable variability in conceptualization and assessment of musical ability and musical training across studies: Some researchers measure perceptual skills with a standardized test, while others assess self-reports of musician status or use a musical intervention. This may introduce variability in the strength and direction of any observed relationship between music and second-language ability. As such, we examined how the operationalization of musical ability related to music–second-language learning effect sizes. Studies that assessed musical ability with a standardized test (e.g., the Profile of Music Perception Skills, or PROMS; Law & Zentner, 2012) or an experimenter-made test were classified as musical *ability* studies. Studies that divided participants into musicians and nonmusicians, asked for participants' self-reported musicianship, or used years of musical experience as a continuous variable were classified as musical *training* studies. Studies that

provided musical training as an intervention (e.g., music lessons versus a control group) were classified as *intervention* studies. However, the very limited number of intervention studies forced this to be treated as a two-level factor (ability/training) for the moderator analyses.

3. Language Measure (Phonology or Non-Phonology): As noted above, work on second-language acquisition in adult learners has suggested particular variability in phonological abilities and morphosyntactic abilities (compared to lexical knowledge; e.g., Flege et al., 1999), and so any relationship with musical ability may differ depending on which aspect of the second language is assessed. Note, however, that most of the literature has focused on second-language phonology (80% of the effects assessed here); thus, we classified effects as either measuring aspects of participants' abilities in second-language *phonology* (80% of effects) or in non-phonological aspects of second-language abilities.

4. Language Modality (Perception or Production): The perception of a second language (e.g., discriminating between phonemes) and production of a second language (e.g., repeating/imitating phonemes) may rely on different abilities. For example, an individual may accurately discriminate between second-language sounds but be unable to produce those sounds in a native-like way. As such, we used a two-level moderator that described if the second-language measure was about language perception or language production. Studies that assessed participants' ability to discriminate between phonemes or discriminate/categorize aspects of second language were classified as *perception* studies and work that assessed participants' abilities to imitate or produce aspects of a second language were classified as *production* studies.

5. Tone/Non-tone Language (Tone or Non-tone): Some work has suggested specific links between musical (pitch) abilities and the learning of languages that use pitch contrastively (i.e., tone languages; e.g., Wong & Perrachione, 2007). Thus, we classified effects as involving a second language (or languages) that was a *tone* language (e.g., Mandarin) or a *non-tone* language (e.g., English). Note that this distinction is only relevant for studies measuring second-language phonology.

a. We identified five studies (with 28 effect sizes) that involved "pitch accent" languages, which use pitch contrastively but to a lesser degree than tone languages (specifically Swedish, Norwegian, and Japanese: Dworkis, 2012; Kempe et al., 2015; Kempe et al., 2012; Nakata, 2002; Narzikul, 2015). This sample was too small to include as a moderator and, because some have argued that pitch accent may not be a coherent category (e.g., Hyman, 2009), we grouped these with the *tone* languages.

b. We also coded whether participants' *first* language was a tone or non-tone language; however, we identified only two studies that indicated that participants' first language was a tone language (Pei et al., 2016; Yang et al., 2014). This sample was also too small to analyze as a moderator.

6. Language Learning Outcome (Proficiency, Novel, or Learning Success): Studies differed in the type of language outcome they investigated. Some examined mastery of the second language, typically based on standardized language proficiency tests; some assessed participants' initial language learning abilities by focusing on the discrimination and/or repetition of novel phonemes; and some measured participants' learning over somewhat longer timeframes (i.e., improvement from pre to post-test, following an intervening language class or experimental teaching intervention). Such effects were classified as assessing *proficiency*, *novel learning*, and *learning success* outcomes, respectively.

Screening & Data Extraction

We conducted abstract and full text screening procedures. First, we curated a list of articles that could match our inclusion criteria from our initial search syntax. Then, two co-authors screened abstracts and full texts to determine if they matched our inclusion criteria. If articles failed one or more inclusion criteria, they were excluded from the subsequent review processes after being checked by another author to ensure accuracy. All coding and exclusion disagreements were resolved by discussion.

We extracted study-level information such as participant age (children or adults), number of participants, participants' native language, the second language assessed, and publication status

(published or unpublished studies). We extracted characteristics on musical ability—specifically how musical ability was operationalized and/or how researchers defined musicians and nonmusicians (if reported). We extracted characteristics on language measures—specifically whether they were perception or production measures and whether they focused on second-language phonology or lexical/syntax measures. Finally, we extracted any relevant effect sizes that quantified a relationship between musical ability and a second-language outcome. If experiments included subcomponents of musical batteries (e.g., pitch perception and rhythm perception for PROMS) and total scores, only subcomponents were included in final analyses to avoid double counting.

Results and Discussion

Main Analyses

To account for multiple correlated effects, we performed a robust variance estimation multivariate meta-analysis using the *robumeta* package (version 2.0; Fisher et al., 2017) in R (version 4.1.2; R Core Team, 2021) with random effects of study where "study" refers to results from one experiment. Each effect size was reported as a Fisher's z-score. Outlier analyses assessed each effect size's studentized residuals, DFITS values, and Cook's distance, which did not indicate any potential outliers or influential data points. Therefore, no data points were removed.

The overall meta-analysis showed a positive, significant relationship between musical ability and second-language learning, $z = .35$, se $= .03$, $p < .0001$, 95% Confidence Interval [.28; .40]. See Figure 2 for a corresponding forest plot (for visualization purposes, this plot shows data aggregated by study; a plot including all 191 effect sizes can be found at https://osf.io/4yzpq/). This analysis shows that past work on musical ability and training (broadly defined) and second-language learning does reveal a positive relationship overall (cf. Chobert & Besson, 2013; Zeromskaite 2014).

However, meta-analytic estimates can be strongly influenced by publication bias. For example, if studies with non-significant or negative effects are less likely to be published (the "file-drawer problem"), this leads to an overrepresentation of larger effect sizes which can lead to an overestimation of, and

potentially incorrect conclusions about, the meta-analytic effect (Maier et al., 2022; Polanin et al., 2016).

Therefore, the next section details how we assessed, corrected, and adjusted for possible publication bias.

Figure 2:

Forest plot of musical ability/training on second-language learning
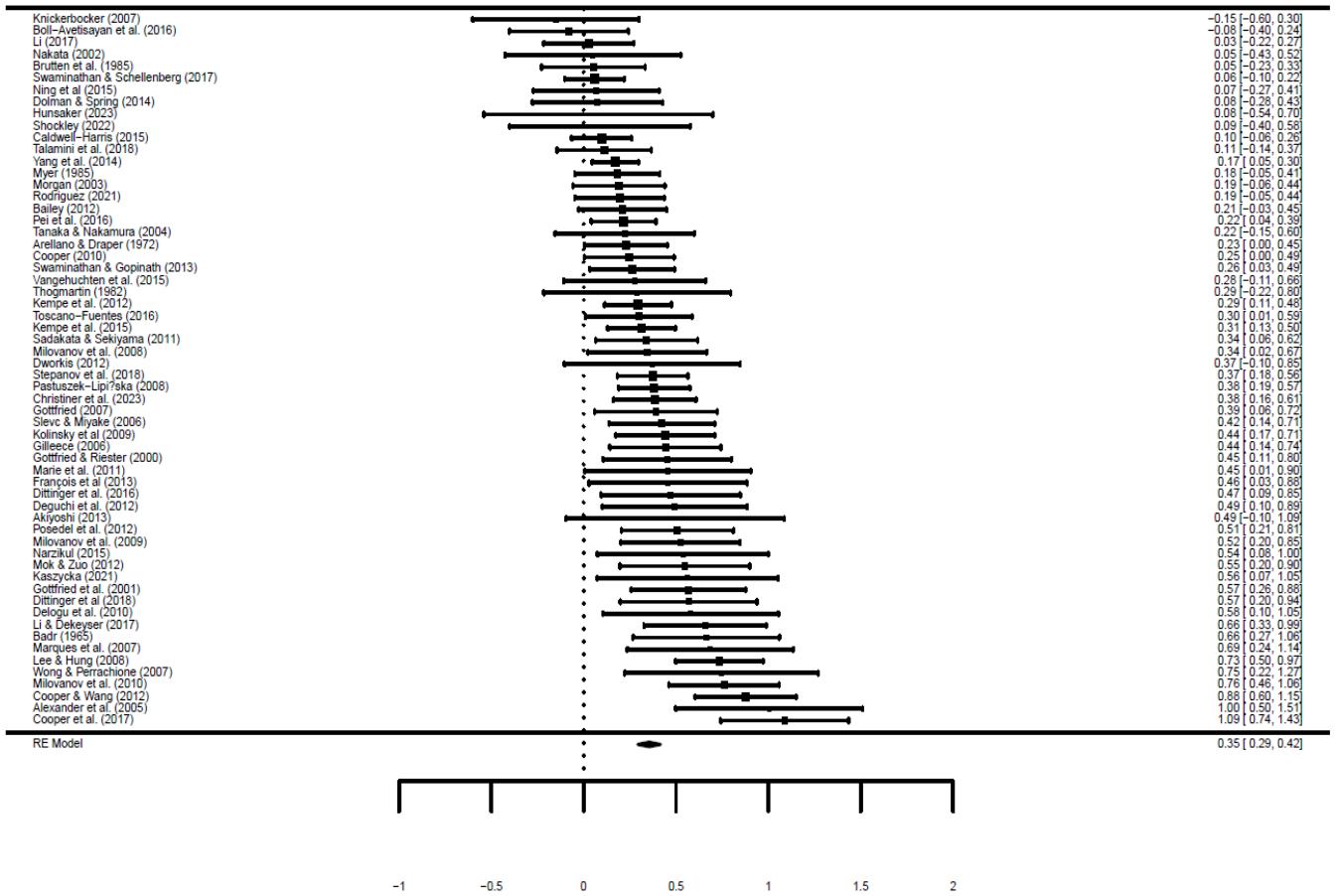
Note. For illustrative purposes, we plotted each experiment's average effect size (Fisher z-score) following guidelines from Hedges et al. (2010) and Link et al. (2014) and ordered them from smallest to largest. Each dot differs in size based on study weight; larger dots indicate larger sample sizes while smaller dots indicate smaller sample sizes. Confidence intervals are given in brackets.

Publication Bias

We relied on three methods to detect and quantify publication bias in our meta-analysis. First, we assessed publication status as a moderator, given that unpublished findings (e.g., unpublished theses and dissertations) are presumably less affected by publication bias than published work. Specifically, we compared 68 effect sizes from the 15 unpublished dissertations/theses identified our literature search to the effect sizes from the remaining 45 published papers (123 effect sizes); these effects did not differ significantly ($z = .26$ and $z = .37$ for unpublished and published studies respectively; $F(1,58) = 3.01$, $p = .08$).

Second, we created a funnel plot (Figure 3) which plots each effect size against its standard error and illustrates the overall meta-analytic effect with a dotted line. In the absence of publication bias, variation around the overall meta-analytic estimate should (a) be symmetrically distributed around the meta-analytic effect and (b) decrease with increasing precision (indexed as decreasing standard error), thus generally occupying the "funnel" shape in white. Visual inspection of the funnel plot suggested some degree of asymmetry, such that the more precise estimates (at the top) appear to be shifted left from the overall meta-analytic effect, and some imprecise estimates (at the bottom) are positive and relatively large. Both of these patterns would be consistent with publication bias. To complement this visual inspection of the funnel plot, we assessed for asymmetry with an Egger's regression test (Egger et al., 1997), which also indicated some funnel plot asymmetry ($t = 5.58$, $p < .0001$, 95% CI [.17 .36]). However, note that an asymmetric funnel plot is not necessarily evidence of publication bias (e.g., Pustejovsky & Rogers, 2019). And even if publication bias exists, this does not necessarily indicate a nonexistent effect, but rather might reflect an inflated estimate of an effect that does, in fact, exist.

Figure 3

Funnel plot of all effect sizes
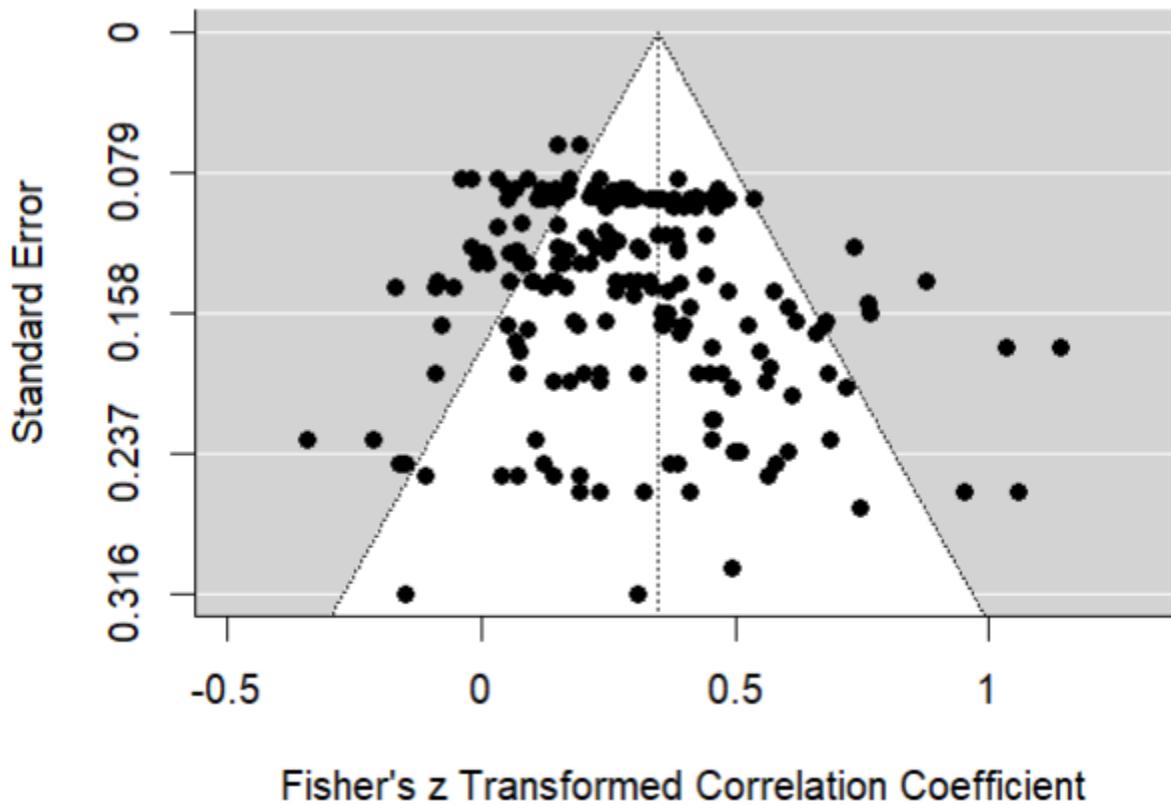
Note. All effect sizes (black dots) plotted against their standard errors. In the absence of publication bias, points should be symmetrically distributed around the meta-analytic effect (dotted line) with variability increasing as precision (as indexed by standard errors) decreases.

Third, we more formally quantified and adjusted for the extent of publication bias by relying on PET-PEESE—the *precision-effect test* and *precision-effect estimate with standard errors* (Stanley & Doucouliagos, 2014). This approach adjusts the relationship between effect sizes and standard errors with two steps: The PET model first predicts effect sizes with their standard errors with a weighted linear regression, while the PEESE model predicts the effect sizes with squared standard errors and then adjusts the overall meta-analytic effect for publication bias (Bartoš et al., 2022). If the PET model is statistically

14

significant, the PEESE test can be interpreted as the meta-analytic effect adjusted for publication bias. The PET test (Table 1) was indeed statistically significant, and accounting for this publication bias yielded an adjusted effect size of $z = .21$, which is notably smaller than the uncorrected meta-analytic effect of $z = .35$ but still indicates a statistically significant positive relationship between musical ability and second-language learning.

Table 1

PET-PEESE Mean Estimates

	Estimate	Standard Error	t	df	p	95 % Confidence Interval	
						Lower	Upper
PET	.14	.03	3.57	187	< .001	.06	.22
PEESE	.21	.02	9.92	187	< .001	.17	.25

Note. Tests conducted in JASP (v. 0.16.2.0; JASP Team, 2023) following guidelines from Bartoš and colleagues (2022).

Taken together, these analyses suggest that publication bias is indeed an issue present within this literature, but it appears to inflate, rather than create, the measured relationship between musical ability and second-language learning. Given this evidence that there is indeed a (modest) relationship between musical and second-language abilities, we assessed how this relationship varies across various potential moderators.

Moderator Analyses

Although the effects making up this meta-analysis all address some aspect(s) of musical and second-language abilities, they come from studies that vary considerably in terms of both participant

populations and methodologies. Therefore, we ran six robust multivariate meta-regressions to assess moderators of interest (Fisher et al., 2017). Table 2 contains a summary of all moderator findings with relevant statistics.

1. Age (Children or Adults): Given discussions of critical and sensitive periods for second-language acquisition (see Flege et al., 1999 for discussion), we investigated if relationships between musical ability and second-language learning differed for adults (here: aged 16 and above) and children (aged 15 and below). Descriptively, children showed smaller effects ($z = .31$) than adults ($z = .35$); however, this comparison was not statically significant ($F(1,58) = .51, p = .47$) Therefore, the relationship between musical ability and second-language learning does not appear to change across age of learning in this sample, although the lack of a moderator effect may be driven by the need to collapse age into a dichotomous category for this analysis rather than investigating age as a continuous variable. Nevertheless, this null result was somewhat surprising given the well-documented advantage in second-language ability for younger learners (e.g., Birdsong & Molis, 2001; Johnson & Newport, 1989), which indicates that younger learners should have less variability in second-language outcomes overall (and thus less variability to be predicted by individual difference factors like musical ability). For example, one might have predicted that adult learners with relatively poor language learning abilities could rely on their musical abilities as a kind of compensatory mechanism, in which case music–second-language relationships should be stronger in adults compared to younger learners (cf. Roman-Caballero et al., 2018 in the case of music practice and age-related cognitive decline); however, we see no evidence to support this idea.

2. Musical Measure (Ability, Training, or Intervention): Typical investigations of music and second-language relationships measure musical ability based on participants' ability to discriminate between minimal musical pairs (e.g., Law & Zentner, 2012; Wallentin et al., 2010), or based on participants' self-reported status as a (non)musician or total years of musical training/experience. More rarely, investigations employ a musical intervention, where listeners are given musical training and are compared to a control group. We thus assessed whether music and second-language relationships varied

by the type of musical measure. Because there were very few musical intervention effect sizes in our sample (n = 4) compared to musical ability (n = 134) and musical training studies (n = 51), our moderator analysis only statistically compared effect sizes for musical ability and musical training studies. There was no significant difference between the effect size in studies measuring musical ability (z = .37) and those measuring musical training (z = .28; F(1, 55) = 3.91, p = .05). Note that, descriptively, the four intervention studies did show a notably larger effect: z = .63. Of course, performance on musical ability tasks is likely correlated with musical experience (Kunert et al., 2016; Wesseldijk et al., 2021), so different relationships of these factors on second-language acquisition may be difficult to observe; nevertheless, these data suggest that second-language learning is similarly related to measures of musical abilities and musical experience.

3. Language Measure (Phonology or Syntax/Semantics): Sound is perhaps the most obvious connection between musical ability and second-language learning, and indeed a majority (80%) of the effect sizes in our sample focus on speech processing (i.e., perception or production of second-language sounds/phonology). However, music and language may be related across multiple levels of representation (Patel, 2008; Slevc, 2012), and some work in our sample assessed other aspects of second-language abilities (e.g., lexical or syntactic). Here, the moderator effect was significant: While both studies involving second-language phonology and studies assessing other aspects of second-language processing yielded significant relationships with musical ability, effects involving second-language phonology were significantly larger (z = .37) than effects involving other aspects of second-language abilities (z = .21; F(1,58) = 42.72, p < .0001).

The finding that musical ability is more robustly related to learning phonology compared to non-phonological aspects of second languages may be unsurprising—simply suggesting that this relationship largely reflects shared auditory processing skills. This could, however, also reflect the lack of diversity in the domains of second-language proficiency assessed in these studies: There were considerably fewer effects assessing non-phonological outcomes (n = 36) compared to phonological (n = 153) aspects of second-language performance.

4. Language Modality (Perception or Production): Accurate speech sound discrimination does not necessarily equate to native-like production (Baese-Berk, 2019), suggesting that musical relationships with perceptual and productive aspects of a second language might differ. Therefore, we assessed if there were differences between studies assessing second-language perception and those assessing second-language production. We observed no significant difference in this moderating effect: Musical ability was similarly related to perceptual ($z = .36$) and productive ($z = .30$) aspects of second-language learning ($F(1, 58) = 2.83, p = .09$).

5. Language Tone Type (Tone or Non-tone): Pitch plays a lexically contrastive role in tone languages which poses difficulties for non-native listeners and speakers (Cooper & Wang, 2012; Wong & Perrachione, 2007). Because pitch plays a central role in musical processing, it is plausible that musical ability and training may relate especially to the ability to perceive, produce, and categorize lexical tone. Therefore, we compared studies of musical relationships with the learning of tone and non-tone second languages. Typically, studies on tone language learning involve categorizing lexical tones (e.g., Dittinger et al., 2018), discriminating between tonal contrasts (e.g., Cooper et al., 2017), or producing various lexical tones (e.g., Li, 2017, Li & DeKeyser, 2017). Indeed, studies investigating tonal aspects of a second language yielded significantly a larger effect size ($z = .54$) than studies measuring learning in non-tonal aspects of a second language ($z = .28$; $F(1,50) = 8.90, p = .004$).

6. Language Learning Outcome (Novel, Learning Success, Proficiency): The studies examined here operationalized second-language ability in multiple ways. Some studies assessed initial abilities/aptitudes with aspects of an unfamiliar second language, for example focusing on the discrimination or repetition of novel second-language phonemes (here, *Novel* studies). Other studies assessed the effectiveness of second-language learning over a short duration, for example by comparing scores on some measure(s) before and after a short second-language learning intervention (here, *Learning Success* studies). And yet other studies assess proficiency in an already-learned second language, typically via performance on some kind of standardized test (here, *Proficiency* studies). Each of these learning assessments could plausibly have quite different relationships with musical ability; therefore, we

18

performed a moderator analysis that assessed how the relationship between musical ability and second-language learning varied based on different second-language outcomes (novel, learning success, or proficiency). Indeed, effect sizes differed significantly between these three groups ($F(2,57) = 4.14$, $p = .02$): Numerically, studies measuring short-term learning yielded the largest effect size overall ($z = .43$) compared to studies examining novel learning ($z = .36$) and studies of eventual proficiency ($z = .22$). Three follow-up robust variance estimations meta-analyses comparing these groups revealed no significant differences between learning success and novel studies ($F(1, 44) = .59$, $p = .44$). Instead, proficiency studies had significantly smaller effects than both learning success studies ($F(1, 22) = 5.21$, $p = .03$) and novel studies ($F(1, 48) = 5.43$, $p = .02$). In summary, this suggests that musical abilities are more strongly related to the immediate (or relatively short-term) perception/production of second-language sounds rather than overall proficiency in the second language.

Table 2

Moderators with analyses (not corrected for multiple comparisons)

Moderator	No. of Effect Sizes	% of sample	z	F(df)	p
Age				$F(1,58) = .51$	.47
Adults	126	67	.35		
Children	63	33	.31		
Musical Measure				$F(1,55) = 3.98_a$	.05
Musical Ability	134	70	.37		
Musical Training (Self Report)	51	28	.28		
Musical Intervention$_b$	4	2	.68		
Language Measure				$F(1,58) = 42.72$	.001
Phonology	153	80	.29		
Syntax/Semantics	36	19	.17		
Language Modality				$F(1,58) = 2.83_c$	.09

Perception	121	65	.36		
Production	64	35	.30		
Tone Language$_d$				$F(1,49) = 8.90$	.004
Tone	26	17	.54		
Non-Tone	127	83	.28		
Language Learning Outcome				$F(2,57) = 4.14$	.02
Novel	108	56	.28		
Learning Success	22	12	.37		
Proficiency	59	32	.19		

Notes.

a. This test does not include the four intervention effect sizes because of limited sample size, the sample for this comparison is reduced to 56 studies.

b. These four intervention effect sizes are not included in the moderator analysis but are included in the overall meta-analysis.

c. This test does not include studies classified as "other" which measured participants' performance on a multiple-choice test of second-language ability.

d. This test only includes the 153 effect sizes involving aspects of second-language phonology

Summary of Moderating Effects: The relationship between musical ability and second-language ability varied across three of the six moderators we assessed: Studies found stronger relationships between musical and second-language ability for phonological compared to lexical/syntactic aspects of second-language learning, for learning of tone compared to non-tone languages, and for immediate perception/production and short-term learning compared to assessed proficiency in a second language. In contrast, relationships between musical ability and second-language learning did not appear to vary as a function of age (in children vs. adult learners), the modality of the language measure (perception vs. production), or the type of musical measure (ability vs. training).

Of course, many other factors might moderate these observed music and second-language relationships, and many of these cannot be easily investigated in group comparisons. However, it is possible to measure and statistically control for some of these individual differences. These sorts of partial

correlation and multiple regression studies can be problematic to include in meta-analyses (e.g., Aloe, 2015), so, in the following section, we detail a narrative synthesis of studies that have controlled for various factors to assess if musical ability can account for unique variance in second-language learning.

Partial Effects (narrative synthesis)

The analyses above focus on studies assessing zero-order correlations between musical and second-language ability. Some studies, however, have investigated this relationship while statistically controlling for other potentially confounding individual differences such as working memory, prior language experience, and intelligence, with the aim of assessing whether musical ability can explain *unique* variance in second-language learning. These types of effects were not included in the meta-analysis above because such partial correlations can have different mathematical properties than bivariate correlations and because correlations adjusted for different covariates are estimating different population parameters and so have different interpretations (Aloe, 2015; Aloe & Thompson, 2013). In addition, if covariates suffer from measurement error or represent an inappropriate causal structure, statistical control can introduce considerable bias (Kahneman, 1965; Wysocki et al., 2022).[3]

Nevertheless, these partial effect sizes certainly play an important role in our understanding of music and second-language relationships, and so here we summarize existing studies of musical ability and second-language learning that reported effects after controlling for various other factors. We followed the same inclusion criteria as our main analysis (see Method) which yielded 12 independent articles reporting 27 partial effect sizes from 643 participants (note that zero-order correlations from some of these same datasets are included in the analysis above). The average sample size in these studies was larger than our main analysis (average N = 72.8 in these 12 studies compared to an average N = 57.7 for the 60 studies above), and these studies primarily assessed musical and second-language learning abilities in adult participants.

---

[3] It should be noted that the inclusion of both bivariate and partial correlations in meta-analyses may be appropriate in some circumstances (see, e.g., Furuya-Kanamori & Doi, 2016); however, the heterogeneity of covariates in this literature suggest that this is not such a case.

Most commonly, studies have controlled for individual differences in prior language experience (e.g., exposure to the second language or starting second-language proficiency). Researchers have also controlled for individual differences in working and/or short-term memory (Bowles et al., 2016; Slevc & Miyake, 2006; Swaminathan & Schellenberg, 2017), intelligence (Badr, 1965, Swaminathan & Schellenberg, 2017; Schellenberg et al., 2023), motivation (Badr, 1965), and experimental context (perception or production training; Li & DeKeyser, 2017). Table 3 lists each study, their second-language and music measures, what they controlled for, and their corresponding partial effect sizes.

In contrast to the clear relationship indicated by the overall meta-analytic results reported above, just over half (53.8%) of these partial effect sizes were statistically significant after controlling for other factors. Nearly all of these positive relationships were observed for studies measuring second-language phonology, although one study (Badr, 1965) found a significant partial correlation between musical ability and second-language vocabulary and comprehension. There was no obvious pattern in the factors controlled for in studies that yielded significant vs. non-significant partial music–language effects. Of course, more than half of these partial effects were significant, and statistical control can artificially inflate both Type I and Type II error (see, e.g., Schisterman, Cole, & Platt, 2009; Westfall & Yarkoni, 2016), so it is not yet clear what to conclude from this small body of work. Nevertheless, this result does underscore the importance of considering other, possibly confounding, factors.

Table 3

Summary of partial effects

Study	N	Second-language Measure(s) (outcome)	Age (Adult/ Child)	Music Measure(s) (predictor)	Partialed Out	Partial Effect Size
Badr (1965)	28	French vocabulary	Child	Pitch Perception	Intelligence	r =.60*
Badr (1965)	28	French vocabulary	Child	Pitch Perception	Motivation	r =.61*
Badr (1965)	28	French comprehension	Child	Pitch Perception	Intelligence	r =.55*

Study	N	Outcome	Age	Type	Predictors	Effect
Badr (1965)	28	French comprehension	Child	Pitch Perception	Motivation	r =.56*
Bowles et al., 2016	160	Mandarin pseudoword accuracy	Adult	Musical Training	Short-term memory	$\beta = .01*$
Cooper & Wang (2012)	54	Post-training Cantonese tone identification	Adult	Musical Training	Pre-training Cantonese tone identification	r = .39
Gotz et al. (2023)	36	Thai speech perception	Adult	Musical Training	Rhythm perception	$\beta = .06*$
Gotz et al. (2023)	36	Thai speech production	Adult	Musical Training	Rhythm perception	$\beta = .25*$
Jacobsen (1974) - Male participants	10	Japanese proficiency	Adult	Musical Training	Prior language exposure/experience, arrival age	$\beta = .02$
Jacobsen (1974) - Female participants	10	Japanese proficiency	Adult	Musical Training	Prior language exposure/experience, Arrival age	$\beta = -.03$
Jekiel & Malarski (2021)	50	English accent production and phonological awareness	Adult	Musical Training	Starting English proficiency, Pitch/Rhythm perception, Gender	$\beta = -.002*$
Li & DeKeyser (2017)	38	Mandarin tone production	Adult	Pitch Perception	Treatment group (perception/production training)	$\eta_p^2 = .13_a*$
Li & DeKeyser (2017)	38	Mandarin tone production	Adult	Pitch Perception	Treatment Group (perception/production training)	$\eta_p^2 =.05$
Mokari & Werner (2018)	38	English phoneme perception	Adult	Pitch Perception	Starting second-language proficiency	r = .29
Mokari & Werner (2018)	38	English phoneme production	Adult	Pitch Perception	Starting second-language proficiency	r = .18
Schellenberg et al., 2023	154	Self rated English proficiency	Adult	Pitch Perception	Gender, Education, Cognitive Ability	r = .06

Schellenberg et al., 2023	154	Self rated English proficiency	Adult	Rhythm Perception	Gender, Education, Cognitive Ability	r = .02
Schellenberg et al., 2023	154	Self rated English proficiency	Adult	Musical Training	Gender, Education, Cognitive Ability	r = .05
Schellenberg et al., 2023	154	Self rated English proficiency	Adult	Singing Ability	Gender, Education, Cognitive Ability	r = .06
Slevc & Miyake (2006)	50	English receptive phonology	Adult	Aggregate of multiple tasks	Age of Arrival, Length of Residence in US, Language Use/ Exposure, Phonological Short-Term Memory	β = .37*
Slevc & Miyake (2006)	50	English productive phonology	Adult	Aggregate of multiple tasks	Age of Arrival, Length of Residence in US, Language Use/ Exposure, Phonological Short-Term Memory	β = .30*
Slevc & Miyake (2006)	50	English grammaticality judgements	Adult	Aggregate of multiple tasks	Age of Arrival, Length of Residence in US, Language Use/ Exposure, Phonological Short-Term Memory	β =.13
Slevc & Miyake (2006)	50	English vocabulary	Adult	Aggregate of multiple tasks	Age of Arrival, Length of Residence in US, Language Use/ Exposure, Phonological Short-Term Memory	β =.04
Swaminathan & Schellenberg (2017)	151	Zulu phoneme identification	Adult	Rhythm Perception	Short-Term Memory Working Memory	r = .25*
Swaminathan &	151	Zulu phoneme identification	Adult	Rhythm Perception	Short-Term Memory, Working Memory,	r = .26*

Schellenberg (2017)					Music Training, Nonverbal Intelligence	
Tabori (2022)	95	Mandarin tone Identification	Adult	Musical Training	Native Language_b	2.47 Odds ratio *

Note. Asterisk (*) indicates significance at $p < .05$ level.
a This value reflects a combined score from three tasks.
b Participants were either Spanish-English bilinguals, Bantu-English bilinguals, or Vietnamese-English bilinguals.

General Discussion & Conclusion

Adult second-language learning success exhibits incredible variability: Some struggle to discriminate between and produce non-native sounds at their desired level, while others reach native-like proficiency (Flege et al., 1999). There are surely many individual differences that contribute to this variability (see, e.g., Li et al., 2023), one of which may be individual differences in musical ability. This is sensible given relationships between musical ability/training and native-language speech processing (e.g., Gordon et al., 2015; Nayak et al., 2022; Neves et al., 2022). However, while many studies have investigated links between musical ability/training and second-language learning, the robustness and consistency of such a relationship has remained unclear (e.g., Mokari & Werner, 2018; Schellenberg et al., 2023). Here, we rely on meta-analytic methods to better understand if, and how strongly, musical ability is related to successful second-language learning.

We synthesized 191 effect sizes across 60 independent articles with a robust variance estimation multivariate meta-analysis. The overall size of the meta-analytic effect was small to moderate according to Cohen's (1992) standards ($z = .35$); furthermore, this effect appears to be inflated due to publication bias. After bias correction methods (PET-PEESE), the meta-analytic effect remained statistically significant, but considerably smaller ($z = .21$): a small effect per Cohen's standards (1992). Thus, we conclude that the effect reflected in the literature is inflated, but not created. This effect size fits with similarly small meta-analytic effects observed for other types of music–language relationships (e.g., Gordon et al., 2015; Neves et al., 2022). While the overall effect at its adjusted size is small, it nevertheless supports a long-speculated relationship between musical ability and second-language

learning and highlights how connections between musical and language processing extend beyond the native language (Besson et al., 2011; Chobert & Besson, 2013; Kraus & Slater, 2015; Patel, 2014; among others). This relationship may also have pedagogical implications for music teaching and second-language learning; for example, these findings fit well with the idea that including music in classrooms or participating in musical activities could assist in the perception and production of non-native sounds (see Amin & Soh, 2020, for a review of music in classroom use). In summary, the present meta-analysis provides evidence that musical ability is indeed positively, albeit modestly, related to second-language learning outcomes across a variety of participant populations and methodologies.

Despite the depth and breadth of the literature examined, the overwhelming percentage of effect sizes assessed here were correlational; thus, this meta-analysis cannot confirm a causal relationship or offer support for "transfer" between music and second-language learning. However, note that the four effect sizes from the three intervention studies included (Cooper & Wang, 2012; Francois et al., 2013; Wong & Perrachione, 2007) found larger effects than musical ability and musical training studies, offering suggestive evidence for some type of transfer. This is plausible given that both music and language are mediated through experience and give rise to training-specific plasticity, likely specific to auditory learning (Herholz & Zatorre, 2012; White et al., 2013). Regardless, the observed relationship is interesting even if it does not reflect a causal effect. Music and language may be the best demonstrations of our ability to perceive and process complex auditory sequences (Patel, 2014). Better understanding music and language relationships could help us understand the underlying mechanisms that give rise to our ability to integrate and extract meaning from sound.

Given the variability in measures and methods across the included studies, we assessed how musical ability and second-language relationships varied across six moderators. Three moderator comparisons were statistically significant: (1) studies measuring second-language phonology yielded larger effect sizes than studies measuring syntax/semantics, (2) studies assessing learning or discrimination of second-language lexical tones yielded larger effect sizes than studies measuring non-tonal aspects, and (3) studies measuring learning success (i.e., change from pre- to post-test) and novel

26

second-language learning yielded larger effect sizes than studies of achieved second-language proficiency. Such differences support a stronger link between musical ability and the processing of second-language sounds (especially pitch) compared to non-auditory and non-pitch aspects of a second language. Critically, however, this link may be limited to (or at least stronger for) the initial perceptive and productive aspects of second-language learning rather than ultimate attainment of second-language proficiency. Such a notion is likely unsurprising given the many factors that influence real-world second-language learning that are not captured with these experiments. (e.g., motivation; Dörnyei, 1998). One possibility is that more musical individuals might have an advantage during early stages of second-language learning due to more accurate perception and production of second-language sounds—likely scaffolded by better auditory processing. However, these differences appear to dissipate over longer periods of time, resulting in smaller relationships with eventual second-language proficiency. In other words, musical ability might predict the "efficiency" of second-language learning, but be less predictive of whether or not learners ultimately achieve high levels of second-language proficiency (see, e.g., Novoa et al., 1988).

Interestingly, the relationships between musical ability/training and second-language learning did not appear to vary as a function of age (i.e., were no different for studies of adults and of children learners), language modality (studies of language perception vs. language production), or assessment of musical training/experience vs. musical ability. While interpretation of such null effects should be approached with caution, it is possible that musical ability relates similarly to both language perception and language production given that perceiving and producing language and music requires fine-grained auditory precision and motor control (Besson et al., 2011; Patel, 2014). Given that there were no observed differences between studies assessing musical training/experience and studies assessing musical ability (via performance on some kind of musical task), it may be that this relationship is not (solely) experience-based (and note that while musical ability and musical training are, of course, related, they are also dissociable; e.g., Law & Zentner, 2012; Mankel & Bidelman, 2018).

Although not part of the meta-analysis itself, we also summarize several studies estimating musical ability's unique contribution to second-language learning (i.e., partial correlation effect size estimates). This analysis was somewhat consistent with the meta-analytic results: Many (but far from all) of these studies found statistically significant positive relationships even after controlling for various other non-musical factors. These included measures of various cognitive abilities (e.g., intelligence or working memory) and demographic factors (e.g., gender, education level). Although the specific factors assessed differed across studies, there was no obvious difference in the covariates controlled for in studies that did and did not yield significant partial correlations between musical ability and second-language outcomes, suggesting that observed correlations do not reflect a single confounding/underlying factor.

Nevertheless, the covariates listed above are only a subset of the many factors that might confound music–language relationships. For instance, socioeconomic status (SES) was not controlled for across any of these studies, yet SES is known to predict both participation and continued involvement in musical activities (e.g., Corenblum & Marshall, 1998, Corrigall et al., 2013; Kinney, 2010; Swaminathan & Schellenberg, 2018) and success in second-language learning (Ariani & Ghafournia, 2016). Personality factors are also associated with both musical involvement (Corrigall et al., 2013) and second-language learning (Chen et al., 2021). Continued research assessing the contribution of these and other potentially confounding factors is clearly an important enterprise (with the caveat that statistical control brings its own set of complications; Schisterman et al., 2009; Westfall & Yarkoni, 2016).

Aside from our primary conclusions, this meta-analysis suggests that the literature on music and second-language learning suffers from publication bias (the unfortunately common "file drawer problem;" e.g., Franco et al., 2014). This could result from researchers not submitting or reporting null/negative findings, the use of various questionable research practices, and/or the difficulty of publishing null results. Of course, any of these reasons are problematic for our understanding of music–language relationships. While the current analysis suggests that publication bias inflates (rather than creates) the overall effect observed here, these statistical corrections are far from perfect (e.g., Carter et al., 2019); thus, this finding underscores a need for large-scale, pre-registered replications.

In addition to publication bias, there are several other limitations across the current literature: A majority of these studies assessed college-aged adults' ability to learn an unfamiliar language. Of these effects, an overwhelming portion of effect sizes (80%) assessed the effect of musical ability/training on phonology-based outcomes, and 65% of the studies assessed perceptual abilities within the second language. This overrepresentation of certain study designs calls for more diverse approaches—specifically, the assessment of second-language syntax, semantics, and production—and a need to investigate more diverse participant populations. (In contrast, the literature has used a highly diverse set of tasks and measures to assess the same underlying constructs, which poses difficulties for cross-study comparisons; Elson et al., 2023).

Conclusions and Future Directions

Learning new languages is becoming increasingly important as the world becomes more interconnected; as such, factors contributing to successful second-language acquisition remain important to investigate. The sum of current evidence suggests a small, but reliable, relationship between musical ability and second-language learning, with the important caveat that this literature also appears to suffer from publication bias. These relationships likely reflect some degree of shared perceptual processing (Chandrasekaran & Kraus, 2010; Patel, 2014) and/or shared cognitive processing (such as auditory attention; Besson & Schön, 2001; Moreno, 2009) that facilitates learning of language sounds. Thus, a clearer understanding of the mechanisms that connect music and (second) language could be gained from work systematically assessing different aspects of auditory processing (e.g., spectral processing and fundamental frequency) in music and second-language learning (e.g., Kachlicka et al., 2019) and assessing relative cognitive skills for processing sounds (e.g., auditory attention). In any case, it appears that musical abilities indeed predict, at least to some extent, success in second-language learning, making this a potentially fruitful area for further exploring music–language relationships.

References

Aloe A. M. (2015). Inaccuracy of regression results in replacing bivariate correlations. *Research Synthesis Methods*, *6*(1), 21–27. https://doi.org/10.1002/jrsm.1126

Aloe, A. M., & Thompson, C. G. (2013). The synthesis of partial effect sizes. *Journal of the Society for Social Work and Research*, *4*(4), 390-405. https://doi.org/10.5243/jsswr.2013.24

Amin, N. A. Z. B. M., & Soh, O. K. (2020). A review of research on the effects of music upon second-language acquisition. *Universal Journal of Educational Research*, *8*(1), 191-201. https://doi.org/10.13189/ujer.2020.080124

Ariani, M. G., & Ghafournia, N. (2016). The relationship between socio-economic status, general language learning outcome, and beliefs about language learning. *International Education Studies*, 9(2), 89-98. http://doi.org/10.5539/ies.v9n2p89

Badr, L. (1965). *Musical Aptitude and the Ability to Learn French as a Foreign Language*. American University of Beirut (Lebanon).

Baese-Berk, M.M. (2019). Interactions between speech perception and production during learning of novel phonemic categories. *Attention Perception Psychophysics*, *81*, 981–1005. https://doi.org/10.3758/s13414-019-01725-4

Bartoš, F., Maier, M., Quintana, D. S., & Wagenmakers, E. J. (2022). Adjusting for publication bias in JASP and R: Selection models, PET-PEESE, and robust Bayesian meta-analysis. *Advances in Methods and Practices in Psychological Science*, *5*(3), 25152459221109259. https://doi.org/10.1177/25152459221109259

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences*, *930*(1), 232-258. https://doi.org/10.1111/j.1749-6632.2001.tb05736.x

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*, 94. https://doi.org/10.3389/fpsyg.2011.00094

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language

    acquisition. *Journal of Memory and language*, *44*(2), 235-249.

    https://doi.org/10.1006/jmla.2000.2750

Boll-Avetisyan, N., Bhatara, A., Unger, A., Nazzi, T., & Höhle, B. (2016). Effects of experience with L2

    and music on rhythmic grouping by French listeners. *Bilingualism: Language and Cognition*,

    19(5), 971-986. https://doi.org/10.1017/S1366728915000425

Bowles, A. R., Chang, C. B., & Karuzis, V. P. (2016). Pitch ability as an aptitude for tone learning.

    Language Learning, *66*(4), 774-808. https://doi.org/10.1111/lang.12159

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology:

    A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological*

    Science, *2*(2), 115-144. https://doi.org/10.1177/25152459198471

Chen, X., He, J., Swanson, E., Cai, Z., & Fan, X. (2021). Big five personality traits and second-language

    learning: A meta-analysis of 40 years' research. *Educational Psychology Review*, *34*, 1-37.

    https://doi.org/10.1007/s10648-021-09641-6

Chobert, J., & Besson, M. (2013). Musical expertise and second-language learning. *Brain Sciences*, *3*(2),

    923-940. https://doi.org/10.3390/brainsci3020923

Christiner, M., & Reiterer, S. M. (2015). A Mozart is not a Pavarotti: Singers outperform instrumentalists

    on foreign accent imitation. *Frontiers in Human Neuroscience*, 9, 482.

    https://doi.org/10.3389/fnhum.2015.00482

Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word

    learning. *The Journal of the Acoustical Society of America*, *131*(6), 4756-4769.

    https://doi.org/10.1121/1.4714355

Cooper, A., Wang, Y., & Ashley, R. (2017). Thai rate-varied vowel length perception and the impact of

    musical experience. *Language and Speech*, *60*(1), 65-84. https://doi.org/10.1177/0023830916642

Corenblum, B., & Marshall, E. (1998). The band played on: Predicting students' intentions to continue

   studying music. *Journal of Research in Music Education*, *46*(1), 128–140.

   https://doi.org/10.2307/334576

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality.

   Frontiers in Psychology, *4*, 222. https://doi.org/10.3389/fpsyg.2013.00222

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, *31*(3), 117–

   135. doi:10.1017/S026144480001315X

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple,

   graphical test. *BMJ*, *315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

   Elson, M., Hussey, I., Alsalti, T., & Arslan, R. C. (2023). Psychological measures aren't

   toothbrushes. *Communications Psychology, 1*(1), 25.

Fisher, Z., Tipton, E., Zhipeng, H., & Fisher, M. Z. (2017). Package 'robumeta'.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition.

   Journal of Memory and Language, *41*(1), 78-104. https://doi.org/10.1006/jmla.1999.2638

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the

   file drawer. *Science*, *345*(6203), 1502-1505. https://doi.org/10.1126/science.1255484

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech

   segmentation. *Cerebral Cortex*, *23*(9), 2038-2043. https://doi.org/10.1093/cercor/bhs180

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the

   effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*,

   42(1), 120-133. https://doi.org/10.1002/ejsp.862

Furuya-Kanamori, L., & Doi, S. A. (2016). Angry birds, angry children, and angry meta-analysts: A

   reanalysis. *Perspectives on Psychological Science*, *11*(3), 408-414.

   https://doi.org/10.1177/1745691616635599.

Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second-language

   phonology. *Brain and Language*, *109*(2-3), 55-67. https://doi.org/10.1016/j.bandl.2008.01.005

Gordon, E. E. (1989). Advanced Measures of Music audiation. Chicago: GIA.

Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology*, *6*, 1777. https://doi.org/10.3389/fpsyg.2015.01777

Gottfried, T. L., Staby, A. M., & Ziemer, C. J. (2004). Musical experience and Mandarin tone discrimination and imitation. *The Journal of the Acoustical Society of America*, *115*(5), 2545-2545. https://doi.org/10.1121/1.4783674

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39-65. https://doi.org/10.1002/jrsm.5

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*, *76*(3), 486-502. https://doi.org/10.1016/j.neuron.2012.10.011

Hyman, L. M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences, 31*(2-3), 213-238.

Jansen, N., Harding, E. E., Loerts, H., Başkent, D., & Lowie, W. (2023). The relation between musical abilities and speech prosody perception: A meta-analysis. *Journal of Phonetics, 101,* 101278. https://doi.org/10.1016/j.wocn.2023.101278

JASP Team (2023). JASP (Version 0.17.2)[Computer software].

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second-language learning: The influence of maturational state on the acquisition of English as a second-language. *Cognitive Psychology*, *21*(1), 60-99. https://doi.org/10.1016/0010-0285(89)90003-0

Kachlicka, M., Saito, K., & Tierney, A. (2019). Successful second-language learning is tied to robust domain-general auditory processing and stable neural representation of sound. *Brain and Language*, *192*, 15-24. https://doi.org/10.1016/j.bandl.2019.02.004

Kahneman, D. (1965). Control of spurious association and the reliability of the controlled variable. *Psychological Bulletin, 64*(5), 326–329. https://doi.org/10.1037/h0022529

Kinney, D. W. (2010). Selected nonmusic predictors of urban students' decisions to enroll and persist in

    middle school band programs. *Journal of Research in Music Education*, *57*(4), 334-350.

    https://doi.org/10.1177/002242940935008

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

    *Reviews Neuroscienc*e, *11*(8), 599-605. https://doi.org/10.1038/nrn2882

Kraus, N., & Slater, J. (2015). Music and language: relations and disconnections. *Handbook of Clinical*

    Neurology, *129,* 207-222. https://doi.org/10.1016/B978-0-444-62630-1.00012-3

Kunert, R., Willems, R. M., & Hagoort, P. (2016). An independent psychometric evaluation of the

    PROMS measure of music perception skills. *PLoS One*, *11*(7), e0159103.

    https://doi.org/10.1371/journal.pone.0159103

Law, L. N., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of

    the Profile of Music Perception Skills. *PloS One*, *7*(12), e52508.

    https://doi.org/10.1371/journal.pone.0052508

Lee, C. Y., & Hung, T. H. (2008). Identification of Mandarin tones by English-speaking musicians and

    nonmusicians. *The Journal of the Acoustical Society of America*, *124*(5), 3235-3248. doi:

    https://doi.org/10.1121/1.2990713

Lee, Y. S., Ahn, S., Holt, R. F., & Schellenberg, E. G. (2020). Rhythm and syntax processing in school-

    age children. *Developmental Psychology, 56*(9), 1632–1641. https://doi.org/10.1037/dev0000969

Li, M. (2017). *Temporal distribution of practice and individual differences in the acquisition and*

    retention of L2 Mandarin tonal word production (Doctoral dissertation, University of Maryland,

    College Park).

Li, M., & DeKeyser, R. (2017). Perception practice, production practice, and musical ability in L2

    Mandarin tone-word learning. *Studies in Second-Language Acquisition*, *39*(4), 593.

    https://doi.org/10.1017/S0272263116000358

Li, S., Hiver, P., & Papi, M. (Eds.) (2023). *The Routledge Handbook of second-language Acquisition and*

    Individual Differences. New York: Routledge.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of internal medicine*, *151*(4), W-65. https://doi.org/10.7326/0003-4819-151-4-200908180-00136

Lin, H (n.d). *Effect Size Converter*. https://www.escal.site/

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*, 861-883.https://doi.org/10.3758/s13423-013-0565-2

Loui, P., Wessel, D. L., & Kam, C. L. H. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception*, *27*(5), 377-388. https://doi.org/10.1525/mp.2010.27.5.377

Maier, M., Bartoš, F., & Wagenmakers, E. J. (2022). Robust Bayesian meta-analysis: Addressing publication bias with model-averaging. *Psychological Methods*, *28*(1), 107–122. https://doi.org/10.1037/met0000405

Mankel, K., & Bidelman, G. M. (2018). Inherent auditory skills rather than formal music training shape the neural encoding of speech. *Proceedings of the National Academy of Sciences*, *115*(51), 13129-13134. https://doi.org/10.1073/pnas.181179311

Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences*, *20*(1), 56-60. https://doi.org/10.1016/j.lindif.2009.11.003

Mokari, P. G., & Werner, S. (2018). Perceptual training of second-language vowels: Does musical ability play a role?. *Journal of Psycholinguistic Research*, *47*(1), 95-112. https://doi.org/10.1007/s10936-017-9517-8

Moreno, S. (2009). Can music influence language and cognition?. *Contemporary Music Review*, *28*(3), 329-345. https://doi.org/10.1080/07494460903404410

Morrill, T. H., McAuley, J. D., Dilley, L. C., & Hambrick, D. Z. (2015). Individual differences in the perception of melodic contours and pitch-accent timing in speech: Support for domain-generality

of pitch processing. *Journal of Experimental Psychology: General*, *144*(4), 730. https://doi.org/10.1037/xge0000081

Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., & Ullén, F. (2014). Practice does not make perfect: No causal effect of music practice on music ability. *Psychological Science*, *25*(9), 1795-1803. https://doi.org/10.1177/0956797614541

Narzikul, K. (2015). Investigating second language learning and musical ability: An ERP study. [Undergraduate honors thesis]. Institutional Repository at the University of Pittsburgh: http://d-scholarship.pitt.edu/id/eprint/24896

Nayak, S., Coleman, P. L., Ladányi, E., Nitin, R., Gustavson, D. E., Fisher, S. E., ... & Gordon, R. L. (2022). The musical abilities, pleiotropy, language, and environment (MAPLE) framework for understanding musicality-language links across the lifespan. *Neurobiology of Language*, *3*(4), 615-664. https://doi.org/10.1162/nol_a_00079

Neves, L., Correia, A. I., Castro, S. L., Martins, D., & Lima, C. F. (2022). Does music training enhance auditory and linguistic processing? A systematic review and meta-analysis of behavioral and brain evidence. *Neuroscience & Biobehavioral Reviews*, 104777. https://doi.org/10.1016/j.neubiorev.2022.104777

Ning, L. H., Loucks, T. M., & Shih, C. (2015). The effects of language learning and vocal training on sensorimotor control of lexical tone. *Journal of Phonetics*, *51*, 50-69. https://doi.org/10.1016/j.wocn.2014.12.003

Novoa, L., Fein, D., & Obler, L. K. (1988). Talent in foreign languages: A case study. In L.K. Obler & D. Fein (Eds.), *The exceptional brain: Neuropsychology of talent and special abilities*, 294-302.

Patel, A. D. (2008). *Music, Language, and the Brain*. Oxford University Press.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98-108. https://doi.org/10.1016/j.heares.2013.08.011

Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural

    distinction of speech syllables. *Neuroscience*, *219*, 111-119.

    https://doi.org/10.1016/j.neuroscience.2012.05.042

Pei, Z., Wu, Y., Xiang, X., & Qian, H. (2016). The Effects of musical aptitude and musical training on

    phonological production in foreign languages. *English Language Teaching*, *9*(6), 19-29.

    http://doi.org/10.5539/elt.v9n6p19

Pelofi, C., & Farbood, M. M. (2021). Asymmetry in scales enhances learning of new musical structures.

    Proceedings of the National Academy of Sciences, *118*(31), e2014725118.

    https://doi.org/10.1073/pnas.2014725118

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between

    published and unpublished effect sizes: A meta-review. *Review of Educational Research*, *86*(1),

    207-236. https://doi.org/10.3102/0034654315582067

Pustejovsky, J. E., & Rodgers, M. A. (2019). Testing for funnel plot asymmetry of standardized mean

    differences. *Research Synthesis Methods*, *10*(1), 57-71. https://doi.org/10.1002/jrsm.1332

Ramjattan, V. A. (2023). Accent Reduction as Raciolinguistic Pedagogy. In Rangan, P., Saxena, A.,

    Srinivasan, R. T., and Sundar, P. (Eds.) *Thinking with an Accent: Toward a New Object, Method,*

    and Practice. https://doi.org/10.1525/luminos.148

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for

    Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roman-Caballero, R., Arnedo, M., Trivino, M., & Lupianez, J. (2018). Musical practice as an enhancer of

    cognitive function in healthy aging-A systematic review and meta-analysis. *PLoS One*, *13*(11),

    e0207957. https://doi.org/10.1371/journal.pone.0207957

Schellenberg, E. G. (2015). Music training and speech perception: a gene–environment interaction.

    Annals of the New York Academy of Sciences, *1337*(1), 170-177.

    https://doi.org/10.1111/nyas.12627

Schellenberg, E. G., Correia, A. I., & Lima, C. F. (2023). Is musical expertise associated with self-reported foreign-language ability? *Journal of Experimental Psychology: Human Perception and Performance, 49*(7), 1083–1089. https://doi.org/10.1037/xhp0001116

Schisterman, E. F., Cole, S. R., & Platt, R. W. (2009). Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology, 20*(4), 488-495.

Slevc, L. R. (2012). Language and music: Sound, structure, and meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(4), 483-492. https://doi.org/10.1002/wcs.1186

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter?. *Psychological Science*, *17*(8), 675-681. https://doi.org/10.1111/j.1467-9280.2006.01765.x

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second-language learning. *Child Development*, 1114-1128, *49*(4), https://doi.org/10.2307/1128751

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60-78. https://doi.org/10.1002/jrsm.1095

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

Swaminathan, S., & Schellenberg, E. G. (2018). Musical competence is predicted by music training, cognitive abilities, and personality. *Nature Scientific Reports*, *8*(1), 1-7. https://doi.org/10.1038/s41598-018-27571-2

Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme perception in a foreign language. *Psychonomic Bulletin & Review*, *24*(6), 1929-1934. https://doi.org/10.3758/s13423-017-1244-5

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

Weber, Andrea, Mirjam Broersma, and Makiko Aoyagi. (2011). Spoken-word recognition in foreign-accented speech by L2 listeners. *Journal of Phonetics*, *39*(4): 479–91. https://doi.org/10.1016/j.wocn.2010.12.004

Wesseldijk, L. W., Mosing, M. A., & Ullén, F. (2021). Why is an early start of training related to musical skills in adulthood? A genetically informative study. *Psychological Science*, *32*(1), 3-13. https://doi.org/10.1177/0956797620959014

Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PloS one, 11*(3), e0152719.

White, E. J., Hutka, S. A., Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Frontiers in Systems Neuroscience*, *7*, 1-10. https://doi.org/10.3389/fnsys.2013.00090

Wilson, D, B. (n.d.). *Effect size calculator.* https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

Wong, P. C., and Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults, *Applied Psycholinguistics*. *28*(4), 565–585. https://doi.org/10.1017/S0142716407070312

Wysocki, A. C., Lawson, K. M., & Rhemtulla, M. (2022). Statistical control requires causal justification. *Advances in Methods and Practices in Psychological Science*, *5*(2), 25152459221095823. https://doi.org/10.1177/25152459221095823

Yang, H., Ma, W., Gong, D., Hu, J. and Yao, D. (2014) 'A longitudinal study on children's music training experience and academic development'. *Scientific Reports*, *4*(1), 5854, 1–7. https://doi.org/10.1038/srep05854

Zendel, B. R., & Alexander, E. J. (2020). Autodidacticism and music: Do self-taught musicians exhibit the same auditory processing advantages as formally trained musicians?. *Frontiers in Neuroscience*, *14*, 1-15. https://doi.org/10.3389/fnins.2020.00752

Zeromskaite, I., 2014. The potential role of music in second-language learning: A review article. *Journal of European Psychology Students*, *5*(3), 78–88. http://doi.org/10.5334/jeps.ci