Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Enjoyment, sex and musical training are associated with duration and passage of time judgements during music listening.

Lígia Silva

ligiaborgessilva@hotmail.com

University of Coimbra

Joanna Witowska

The Maria Grzegorzewska University

Luke Jones

University of Manchester

Michelle Phillips

University of Manchester

Article

Keywords: Time perception, age, sex, musical expertise, duration estimates, passage of time

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11557097/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Authors

Lígia Borges Silva[1]*, Joanna Witowska[2], Luke Jones[3], Michelle Phillips [3,4]

Affiliations

[1] Centre for Interdisciplinary Studies, Faculty of Arts and Humanities, University of Coimbra, Portugal;

[2] The Maria Grzegorzewska University, Poland

[3] University of Manchester, UK

[4] Royal Northern College of Music, Manchester, UK

Ligia Silva - https://orcid.org/0000-0001-5095-1929

Joanna Witowska: https://orcid.org/0000-0002-9063-049X

Luke Jones: https://orcid.org/0000-0002-2202-3276

Michelle Phillips: https://orcid.org/0000-0003-0933-3621

Corresponding author:

Lígia Borges Silva: ligiaborgessilva@hotmail.com

Title: Enjoyment, sex and musical training are associated with duration and passage of time judgements during music listening.

Abstract

Music is inherently temporal, which led multiple studies to focus on how music listening affects time perception. However, little is known about how listeners' individual characteristics affect duration estimates and passage of time judgements while listening to music. We pooled data from seven online experiments on music and time perception (437 adults) to examine the role of age, sex and musical expertise, also considering individual levels of enjoyment while listening. After listening to musical pieces lasting from 16 seconds to c.4.5 minutes, participants estimated perceived duration and rated how fast/slow time had passed. Both retrospective and prospective timing paradigms were investigated. Higher enjoyment was associated with a faster perceived passage of time in both paradigms, and listeners reporting more active engagement with music in daily life enjoyed the pieces more. Male participants gave longer duration estimates than females in both paradigms, although only when durations were reproduced rather than estimated verbally. A greater level of musical training was associated with longer and more accurate prospective estimates, while age seemed not to interfere with time perception. These factors explained some of the variance in duration estimates, and suggest that other listener characteristics may also shape how musical time is experienced.

Keywords: Time perception; age; sex; musical expertise; duration estimates; passage of time

Introduction

Time perception is a complex phenomenon that includes different and seemingly independent temporal judgements: duration estimates (DE); and the sense of how fast/slow time passes, labelled as passage of time (PoT) [1]. Time can be perceived prospectively, when one reflects

on time while experiencing it; or retrospectively, when one estimates the duration and passage of time of a past lived event [2].

Current models of time perception

Regarding experience of duration, there are currently two main types of models derived from empirical evidence, which attempt to represent how duration is perceived in different contexts [3]. For retrospective DE, memory-based models propose that perceived duration depends upon how much information is stored, how it is segmented, and the cognitive effort associated with its processing [4,5]. The more units of information are stored and the greater the processing effort, the longer the interval will be perceived, retrospectively. For prospective DE, internal-clock models describe a neural mechanism, where pulses are constantly being created and accumulated to generate a duration estimate [6,7]. Evidence shows that the rate of internal pulses can be affected by the rate of an external regular stimulus, via an entrainment mechanism [8-10]. The faster the rate of that stimulus, the faster the internal clock, leading to longer prospective estimates of duration. Attention to concurrent tasks can lead to less attention being paid to time. As a result, some pulses may not be accumulated, leading to shorter DE [7]. DE may be studied both in terms of a bias towards over- or underestimation, or of overall accuracy. However, no hypotheses were formulated regarding accuracy, since evidence on accuracy is scarce.

Regarding experience of the passage of time, evidence has suggested that this is consistently affected by experienced emotions [3, 11-13]. Positive emotions and engaging tasks tend to lead to faster perceived passage of time, as opposed to negative emotions and boring tasks, when experienced both prospectively and retrospectively.

Music can be considered a relevant contextual factor for time perception. Music is a temporal stimulus by itself. Music provides regular temporal markers that can influence prospective DE according to internal clock models [14, 15], affords different degrees of segmentation of information which can affect retrospective DE [16, 17], and can also foster different levels of engagement that appear to affect passage of time judgements [18]. Some studies have focused on how different musical parameters may affect time perception, such as tempo, meter, tonality, or timbre, also considering individuals' familiarity with the music [14, 15, 18-22].

In terms of individual differences related to music and time perception, while the role of musical expertise on time perception has already been studied in the context of music listening, to our knowledge, other individual factors such as age and sex have not yet been considered extensively, even though they received some attention in the scope of broader time literature. Since some authors have suggested that time perception may operate differently in musical contexts than in other contexts [23, 24], it is therefore relevant to study how these individual factors influence time perception while listening to music. The main aim of this study is to investigate the role of individual factors on time perception during music listening, across similar musical contexts, while allowing for some lower-level musical variation. Since a relatively large sample is needed in order to effectively investigate these

factors, we pooled datasets coming from different but comparable experiments on music and time perception.

Musical expertise

Musical expertise has been conceived and assessed differently across empirical studies. While older studies conceptualized it mostly via years of formal musical training and a musician/non-musician dichotomy, later studies have employed more advanced understanding and assessment tools such as aural tests, or the Goldsmiths Musical Sophistication Index (Gold-MSI). Musical sophistication is a multi-dimensional construct, including five subscales: Active Engagement (AE), Emotions (Em), Perceptual Abilities (PA), Musical Training (MT) and Singing Abilities (SA) [25].

It is still unclear whether musical expertise affects how time is perceived while listening to music, a fact that may derive from the different ways in which musical expertise is assessed. Many skills associated with music perception and motor synchronization with music can be learned passively through exposure and engagement with music, placing musicians and non-musicians at similar levels [26]. However, musicians appear to have better perceptual sensitivity to timing deviations [27, 28], especially in temporal tasks lasting less than 1second [29]. Compared to non-musicians, musicians' DE seem to be less influenced by tasks including a motor component, such as a concurrent tapping task [29, 30], and by the emotions conveyed in music [31].

According to existing models of time perception, music is expected to shape subjective time via differences in memory, attention, and enjoyment. Regarding memory, musicians can process more musical material, more accurately, and faster than non-musicians [32], and have better working memory for tonal and verbal material than non-musicians [33] Aligned with these findings, Phillips and Cross [34] found that musicians retrospectively estimated the duration of a musical piece as shorter, and more accurately, than non-musicians, concluding that musicians probably spent less effort in music processing and likely structured the music in fewer and more comprehensive structural units. Regarding attention, higher musical expertise is expected to lead to less cognitive effort in processing music and more attention dedicated to tracking time, with prospective duration being perceived as longer. While one study confirms this [35], another study found no such effects [20]. Higher levels of musical expertise may only improve prospective timing for more complex musical stimuli, which would align with differences in musical stimuli complexity between those studies. Regarding enjoyment, participants who have stronger and more frequent emotional responses to music, and who more actively seek music-related activities, are expected to experience more positive experiences with the music and report higher levels of engagement and enjoyment, leading to faster passage of time judgements.

The five Gold-MSI subscales are differently related to each of these factors. The AE and Em subscales assess emotional responses and engagement with musical activities, and may better

predict PoT judgements. The PA and MT subscales correlate more highly with scores from aural tests [25] and relate more closely to Ericsson and Pool's [36] concept of expertise, which focusses on hours of deliberate practice and the ability to operate with mental representations of music. Scores for these subscales may better assess cognitive skills involved in memory and attention, mainly predicting DE.

Taking into account findings from previous research we hypothesized that:

H1 MS. R.D. – Higher scores for the PA and MT subscales will be negatively associated with DE, retrospectively.

H2 MS. P.D. - Higher scores for the PA and MT subscales will be positively associated with DE, prospectively.

H3 MS. R. PoT- Higher scores for the Em and AE subscales will be associated with higher enjoyment of the music and with faster PoT, retrospectively.

H4 MS. P. PoT- Higher scores for the Em and AE subscales will be associated with higher enjoyment of the music and with faster PoT, prospectively.

Age

Although our mental representations of time, and the relevance attributed to it change during our lifetime [37, 38], the evidence is not so clear on how age impacts the perception of duration and passage of time in the range of a number of minutes.

Regarding prospective DE, timing performance does not seem to vary between older and younger adults [38-40]. The existing evidence of a lower and more variable performance in timing tasks appears to be caused not by higher age alone, but instead by the age-related cognitive decline in attention and memory [37, 40, 41]. Evidence also shows that the temporal information processing of older individuals is preserved in the case of active musicians [40]. Furthermore, the differences in timing performances seem to depend on the timing task used (generalization vs bisection tasks), with inconsistent results being found among studies, especially in the sub-second range [38, 39, 42]. While these pieces of evidence disagree on a possible effect of age on timing tasks under one second, they tend to converge towards a potentially null impact of age on prospective DE beyond 15 seconds and specifically in the minute range, the time interval most relevant for music listening experiences. To our knowledge, there is no research specifically looking into age-related differences in prospective DE concerning musical activities, such as listening or performing music.
Most reported age-related changes in time perception refer to retrospective judgements of long durations (e.g. years, decades) [42]. For shorter time intervals (30-120 seconds), the impact of age seems to depend on the task. While older participants provided shorter DE than younger participants in tasks involving text and images [43], the contrary was found for a task of listening to music [34]. Phillips and Cross's [34] experiment is one of the only such studies to focus on music listening. Since age-related effects on time perception seem to be

task-dependent, it is possible that age impacts the perception of musical duration differently than other tasks, especially given the inherently temporal dimension of music.

It is generally agreed that time is felt to be moving faster the older we are. However, this assumption seems to conflate the perceived speed of the PoT with duration, and with the individual's subjective value of a time interval [8]. Although evidence confirms that the last decade of one's life tends to be perceived as passing faster as the individual's age increases, the same does not seem to occur in the case of years, months, or hours [38, 44, 45]. Two studies investigated potential age-related differences in the retrospective perceived speed of the PoT in the hour-minute range, during everyday activities [11, 12]. Both studies found no overall significant differences in PoT related to age, and a correlation between negative emotions and slower judgements of PoT, independent of the participant's ages. Considering this evidence , we do not expect any age-related differences to be observed in participants' judgements about the speed of the PoT. Three hypotheses were formulated regarding age:

H5 age R.D. - Age will be positively associated with retrospective DE.

H6 age P.D. - Age will not directly affect prospective DE.

H7 age PoT – Age will not directly affect retrospective judgements of the PoT.

Sex

This section will focus on differences in time perception related to biological sex. The existing literature is inconclusive on whether there are sex-related differences in time perception, and in what situations these may occur. Regarding retrospective duration judgements, a meta-analysis including 74 experiments concluded that under a retrospective paradigm, females tend to provide retrospective verbal estimates of duration that are about 10% longer than males [46]. Authors attributed these findings to females' better episodic memory, suggesting that females may recall more information from the interval to be timed, an explanation consistent with memory-based models [46, 47].

The existing evidence on prospective DE presents diverging results depending on the methods used. Block et al. [46] concluded that sex did not significantly impact prospective DE. However, this null effect was the result of a combination of positive effects (females providing longer estimates than males) found in experiments using verbal and production methods, and negative effects (males providing longer estimates) found in experiments using reproduction tasks. The reproduction method, when compared to production and verbal estimation methods, appears to lead to fewer sex differences [46]. Method-related differences in prospective duration judgements were more recently confirmed [29]. However, other studies which also employed the production method contradicted findings from Block et al. [46] regarding the production method, finding that females produced target durations shorter than males [48-51]. Some of these studies suggested that females perceived the target duration as being reached sooner than males, and attributed this to a possibly faster internal clock in females [51].

Studies which focus on accuracy instead of the bias towards under- or overestimation suggest some evidence of males being more accurate than females in prospective DE in reproduction and comparison tasks [52], and of females' accuracy being more sensitive to environmental context than those of males [53]. However, higher accuracy for males was mostly observed for tasks with durations below 5 seconds[52, 53], and it remains unclear whether these differences occur for longer tasks of a number of minutes, such as music listening.

It is important to note that none of the above-mentioned research referred to a music listening situation. In relation to the reproduction method used in all of the prospective experiments included in this study, we hypothesize that our results for the prospective condition will replicate conclusions from Block et al.'s [46] meta-analysis, with females reproducing durations shorter than males. Due to the lack of evidence on sex effects on PoT judgements and overall accuracy on duration estimates, no hypothesis wasformed regarding these temporal judgements. Two hypotheses were formulated for the effects of sex on DE:

H8 sex R.D. - Females will retrospectively estimate duration as longer than males.

H9 sex P.D - Females will prospectively estimate duration as shorter than males.

Methods

The dataset

One of the challenges presented by existing literature examining perception of time is that experiments differ, often significantly, in aspects of their methodology, and therefore this limits the comparability of results across studies. The current paper seeks to address this by combining data from 7 different experiments, all undertaken by one or more authors of this paper, each focused on studying the impact of one or two musical parameters on time perception. Although the studies were originally conducted separately, similar methodologies including stimuli, procedures, and measurements made it possible to integrate the datasets for the present analysis. The experiments shared the following features: Stimuli:

- Instrumental music with no lyrics, including only 3 possible instruments: voice, flute and piano.
- Tonal music with regular meter.
- Unfamiliar music, composed specifically for use in empirical research
- Relatively limited duration range: from 16 seconds to 4 minutes

Procedures:

- Conducted online
- Listening to music only, no concurrent task

Measures:

- Passage of time, assessed via Likert scales
- Perceived duration assessed mostly via reproduction method for DE
- Musical sophistication assessed with the Gold-MSI

Some participants took part in more than one experiment, during the same empirical session: 30 in Timbre Exp. 1 and Tonality Exp. 3, 10 in Meter and Tonality Exp. 3, and 43 in Meter and Timbre Exp. 1. In addition, 121 participants contributed to both paradigms: 80 participants of Tempo Exp. 3 completed prospective tasks (Meter, Timbre Exp. 1 or Tonality Exp. 3) after the retrospective task in the same session, and 41 participants of Tonality Exp. 2 completed both paradigms. The retrospective and prospective samples therefore partly overlap. Table 1 describes the different experiments and the data collected in each of them.

Table 1
Number of Participants and Observations per Experiment of Origin

Experiment of origin	Participants	Retro. DE	Retro. PoT	Prosp. DE	Prosp. PoT	Gold-MSI
Tempo Exp. 3	132	128	131	—	—	Reduced
Tonality Exp. 2 [15]	116	63	—	353	—	Reduced
Tempo Exp. 4 (verbal estimates)	161	161	161	—	—	Complete (38 items)
Timbre Exp. 1 [18]	73	—	—	73	73	Reduced
Timbre Exp. 2 [18]	21	—	—	84	84	Not collected
Tonality Exp. 3	41	—	—	39	41	Reduced
Meter	56	—	—	156	156	Reduced
Total	437 (unique)	352	292	705	354	

Note. Counts after removal of outlying trials and of the 2 participants who reported another sex category or did not answer. Prospective cells count observations (trials); several participants contributed up to four trials and took part in more than one experiment. DE = duration estimates; PoT = passage of time; — = not collected. Reduced = Gold-MSI versions with two to four items per subscale; Complete = full 38-item inventory.

The Gold-MSI is used in multiple of the studies here, in its subscale format. The scale was designed to allow reductions of this questionnaire, with scores being calculated as averages per subscale in a dedicated app (https://shiny.gold-msi.org/gmsiscorer/). In this way, the

different formats across experiments are not expected to impact the results significantly, since reduced formats included items from all five subscales.

Some of the studies considered in this dataset also included trials presenting more distinctive musical stimuli. For instance, the tonality study included trials presenting atonal music and irregular, randomly generated rhythms, and the timbre experiments included artificial voices. Those were considered not consistent with the remaining stimuli from other experiments, and those trials were hence excluded from this dataset. This consistency provides more confidence that musical experiences across these different studies were comparable enough to be analyzed together.

Importantly, most DE in these studies were reproduction estimates, where the participant reproduces the perceived duration of the music after listening to it, with a start/stop button. This method was considered to represent differences in subjective duration more accurately than other methods (such as verbal estimation, comparison or production) and is especially suitable for shorter musical stimuli [15,35]. Only Tempo Exp. 4 used verbal estimation instead of the reproduction method. This difference was due to the fact that that particular experiment had relatively longer stimuli (3:31-4:27 minutes, compared with 16-38 s) for which a verbal method is more suitable than the reproduction method [38].

Participants

This dataset includes responses from 439 participants. Participants were recruited via social networks, Prolific and Survey Circle. Participants recruited via Prolific received monetary compensation of around €6.00 per hour. Two participants who reported another sex category besides Female and Male, or did not answer this question, were excluded, leaving 437 participants for the analyses. Table 2 shows descriptive statistics for age, sex and Gold-MSI subscales for the 437 participants included in the analysis.

Table 2

Demographic Characteristics and Gold-MSI Subscale Scores of Participants

Variable	*n*	*M*	*SD*	Min	25th	Mdn	75th	Max
Age (years)	436	36.36	14.42	18.00	24.00	32.00	44.00	78.00
Active Engagement	416	4.15	1.27	1.00	3.33	4.00	5.00	7.00
Perceptual Abilities	416	4.94	1.07	1.00	4.33	5.00	5.67	7.00
Musical Training	416	2.82	1.71	1.00	1.29	2.29	4.00	7.00
Singing Abilities	416	4.01	1.36	1.00	3.00	4.00	5.00	7.00
Emotions	416	5.42	1.21	1.00	4.50	5.50	6.50	7.00
Sex								
Female	251	57.4%						
Male	186	42.6%						

Note. $N = 437$ participants included in the analyses. Gold-MSI subscale scores range 1–7. *M* = mean; *SD* = standard deviation; Min and Max = observed range; 25th, Mdn and 75th = percentiles. For Sex, the *M* column shows the percentage of participants.

Materials

The stimuli consisted of several short musical compositions created by a composer, especially for use in empirical research focused on music and time perception. Table 3 describes the musical characteristics of the stimuli including: name, instrumentation, tempo and duration.

Table 3
Musical Characteristics of Presented Stimuli

Experiment of origin	Instrumentation	Name of Stimulus	Tempo (BPM)	Duration
Tempo Exp. 3	Flute	Melody A	60	35 s
			150	38 s
Tonality Exp. 2 [15]	Flute	Melody A	120	19 s
			90	26 s
		Melody B	120	16 s
			90	21 s
Tempo Exp. 4	Piano	Time to Go (by Matthew Woolhouse)	86	267 s
			98	236 s
			110	211 s
Timbre Exp. 1 [18]	Female voice	Natural melody	90	31 s
Timbre Exp. 2 [18]	Female voice	Happy 1	70	38 s
		Happy 2	100	31 s
		Sad 1	66	39 s
		Sad 2	100	32 s
Tonality Exp. 3	Flute	Melody B - slow	90	21 s
Meter	Female voice + piano	Melody with bass in quarter-notes	120	20 s
		Melody with bass in eighth-notes	120	20 s
		Melody with bass in half-notes	120	20 s

Note. BPM = beats per minute; s = seconds.

To assess musical expertise we used the Gold-MSI inventory, which is comprised of 38 items rated on 7-point scales, distributed across five subscales: Active Engagement (AE, time, effort, and resources devoted to musical activities), Perceptual Abilities (PA, self-assessed musical listening and perceptual skills), Musical Training (MT, formal training, practice), Singing Abilities (SA, singing-related skills), and Emotions (Em, the use of music for emotional purposes). The questionnaire also includes an 18-item General Musical Sophistication factor, derived from items across the five subscales, which was not considered in this study. The original validation study reported good internal consistency across the subscales (Cronbach's α = .789–.900) and high test–retest reliability (r = .857–.972), as well as evidence of convergent validity with established musicality and listening tests [25]. Tempo Exp. 4 employed the complete inventory, while the remaining experiments used reduced

versions with two to four items per subscale, created using the dedicated application (https://shiny.gold-msi.org/).

Procedure

All experiments were conducted online, with stimuli being presented either via a Wix website or LimeSurvey, and data being collected in LimeSurvey or Qualtrics. After providing informed consent, participants were asked to proceed only if they were in a quiet place, and to listen with their eyes closed, using headphones. Instructions asked participants to use the fullscreen on their computer, which hides the clock display in devices (to prevent participants from using the digital clock to help with time estimates). Participants would then listen to each stimulus included in that experiment (one by one, when several were presented), providing temporal judgements for each separately.
In all experiments, right after listening to each stimulus, participants estimated the duration and rated the perceived speed of the passage of time. DE were asked for first, followed by ratings of the passage of time. In experiments including both retrospective and prospective paradigms, stimuli and temporal judgements regarding the retrospective paradigm were always presented first, right after the informed consent and the initial instructions. If a retrospective temporal judgement was included, the experiment instructions mentioned only the focus on music perception, but not on time perception, an important procedure to correctly study retrospective time judgements (participants should not be aware before they begin the task that they will be asked to give an estimate of time, which may lead to them using strategies to track this while listening). After the retrospective judgement was provided, a message fully disclosed the objects of study, and the participant was invited to proceed with the other prospective tasks.

When using the verbal estimation method (exp. Tempo 4), participants were asked to estimate the amount of time that had elapsed, filling in blank spaces for ___ minutes and ___ seconds, in response to the following question: “Intuitively and without looking at your watch or a clock, how long do you think you were listening to the music?”. When using the reproduction method participants were asked to reproduce the perceived duration of the musical stimulus using a start/stop button. Passage of time judgements were elicited using the question “How fast did time pass for you while you were listening to music?”, followed by a Likert scale. Two Likert scales were used: from 1- extremely slow to 7 extremely fast (Tempo Exp. 4); or 1 – very slow, 2 – slow, 3 – normal, 4 – fast, 5 – very fast (for the remaining experiments).

Most experiments (except for the meter experiment) also asked participants about experienced pleasure/enjoyment for each stimulus, via a question paired with a Likert scale. The questions used were: “How much did you enjoy listening to the music?” (1 – I didn’t like it at all, 7 – I liked it a lot, for the exp. Tempo 4), and “How pleasant was the music you just listened to?” (1 – very unpleasant, 2 - unpleasant, 3 - neutral, 4 - pleasant, and 5 - very

pleasant, for all the remaining experiments). After listening to all stimuli, participants indicated age, sex, and answered questions about musical expertise (except for Timbre Exp. 2, where musical expertise was not assessed).

Data analysis

Data preparation

DE were converted into relative estimates (participants' estimates / real duration of that particular stimulus), where a value of 1 would represent an accurate estimate, values under 1 an underestimate, and above 1 an overestimate. Outliers were identified from each experiment's full prospective dataset following the 1.5 IQR rule. Cut-off values were computed on the full original datasets. Relative DE were subsequently z-scored within experiment of origin and paradigm. This makes the verbal and reproduction estimates comparable, and gives every experiment a mean of 0 and a standard deviation of 1. Differences in the overall level of responses between experiments, for instance due to stimulus duration, scale length or estimation method, were therefore removed before the analysis, and experiment of origin was not included as an additional factor in data analysis. The analyses thus examines how individual factors relate to a participant's judgements relative to the other participants in the same experiment. PoT judgements and enjoyment ratings were also z-scored per experiment of origin, to facilitate comparability in Likert scales across the different experiments (5 vs 7-points). Pleasantness and enjoyment questions were considered as equivalent adjectives and experiences, and were therefore combined as the 'enjoyment' variable in the analyses.

Statistical analysis

Retrospective and prospective paradigms were treated separately. The analysis considered four possible combinations of factors: two paradigms (retrospective and prospective) x two temporal judgements (DE and PoT). Some participants listened to more than one musical stimulus. For instance, some participants took part in both Timbre Exp.1 and Tonality Exp. 3. As a result, prospective data could include several observations per participant, while retrospective data contained only one observation per participant. Enjoyment ratings were considered in the analysis as a possible covariate influencing PoT judgements, DE, and accuracy.

Shapiro-Wilk tests indicated no normal distribution of DE and PoT judgements, for both prospective and retrospective conditions. Retrospective temporal judgements were initially analyzed using Spearman correlations (for musical sophistication scores – H1 and H3, for age H5) and Mann-Whitney U tests (for sex – H8). When these tests revealed significant effects for more than one individual factor, a multiple linear regression was calculated including these factors as predictors, to investigate their relative strength on the dependent variable. Our hypotheses regarding prospective temporal judgements were all tested by analyzing results using linear mixed models (LMM). These models can effectively

analyze data with an unequal number of observations per participant (unlike repeated measures ANOVAs) and account for possible individual baseline differences in time perception across participants. First, for each temporal judgement (the dependent variables) a series of LMMs were calculated, each having only one individual factor (age, sex, or Gold-MSI subscale) or enjoyment as a fixed effect or covariate. All models included participants' intercept as a random factor. When significant effects were found for more than one factor, a final model was fitted, testing the inclusion of those factors and comparing AIC values under a maximum likelihood estimation method. In those cases, residuals' distribution was inspected visually and via Shapiro-Wilk tests, and log transformations of the dependent variable were considered as a resource to potentially improve model fit and residuals' distribution. These were preliminarily tested when applicable and reported when they proved beneficial.

In terms of duration, besides using relative estimates (estimates/real duration) to analyze whether individual effects created a directional bias towards under- or overestimates, an extra step in the analysis focused on the accuracy of DE judgements and magnitude of temporal deviation. For this, we calculated the absolute error, which is the positive value of the difference between the relative estimate and real duration: abs(relative DE - 1) [47]. The lower the absolute error, the more accurate the estimate. Absolute error was standardized within experiments. It was studied using an exploratory approach, and considered as a dependent variable, and analyzed in the same way as DE and PoT judgements.

A sensitivity analysis ($\alpha = .05$, two-tailed, power = .80; G*Power [54-55]) showed that the retrospective analyses could detect correlations of $|\rho| \geq .15$ for DE ($n = 352$) and $|\rho| \geq .16$ for PoT judgements ($n = 292$), and sex differences of $d \geq 0.31$ (206 females, 146 males) and $d \geq 0.34$ (171 females, 121 males), respectively (Mann–Whitney test, A.R.E. method). For the prospective LMMs, a conservative estimate based on the number of participants rather than observations indicates sensitivity to effects of $|r| \geq .20$ for DE (201 participants) and $|r| \geq .27$–$.30$ for PoT judgements (87–108 participants). Smaller effects may have gone undetected, particularly for prospective PoT. Five observations from two participants who reported another sex category or did not answer were excluded because these categories were too small for inference.

For Mann–Whitney U tests, effect sizes were reported using rank-biserial correlation (r_rb). For LMMs, both marginal and conditional R^2 were reported. Hypotheses tested with two Gold-MSI subscales (H1–H4) were Holm-corrected for two comparisons, and all significant hypothesis-related results remained significant after correction. Analyses that were not hypothesized (accuracy, and enjoyment as a predictor of DE) are exploratory and were not corrected for multiple comparisons. Residuals of the prospective duration models deviated slightly from normality (Shapiro–Wilk $p < .05$; skewness ≈ 0.2), and those of the absolute-error models were right-skewed. Since log-transforming relative estimates did not improve the residuals and fixed-effect estimates of LMMs are robust to such deviations in samples of this size [56], untransformed variables were kept.

Results

Sample sizes differ between analyses because not all original experiments collected data relating to all measures (Table 1). In this section, only significant findings are reported in full. Values for non-significant findings can be found in Tables 4 and 5, along with the sample size of each test.

Table 4

Summary of Findings Regarding Passage of Time and Enjoyment

Predictor	Retrospective test	Retrospective result	Prospective result (LMM)
A. Predictors of passage of time (PoT) judgements			
Sex	Mann–Whitney U	U = 11,531.5, p = .091, r_{rb} = −.11; n = 292 (171 F, 121 M); males M = −0.104, SD = 0.969; females M = 0.073, SD = 1.021	b = −0.230, SE = 0.146, z = −1.58, p = .115, 95% CI [−0.517, 0.056]; n = 354 obs. (108 participants)
Age	Spearman	ρ = −.017, p = .769, n = 292	b = 0.000, SE = 0.007, z = 0.04, p = .965, 95% CI [−0.013, 0.014]; n = 354 obs. (108 participants)
AE	Spearman	ρ = .197, p < .001, n = 292; Holm-corrected p = .001	b = 0.030, SE = 0.063, z = 0.48, p = .630, 95% CI [−0.094, 0.154]; n = 270 obs. (87 participants)
PA	Spearman	ρ = .001, p = .984, n = 292	b = 0.045, SE = 0.075, z = 0.59, p = .553, 95% CI [−0.103, 0.192]; n = 270 obs. (87 participants)
MT	Spearman	ρ = .014, p = .815, n = 292	b = 0.050, SE = 0.044, z = 1.13, p = .260, 95% CI [−0.037, 0.136]; n = 270 obs. (87 participants)
SA	Spearman	ρ = −.034, p = .567, n = 292	b = −0.072, SE = 0.059, z = −1.22, p = .221, 95% CI [−0.186, 0.043]; n = 270 obs. (87 participants)
Em	Spearman	ρ = .040, p = .494, n = 292	b = 0.051, SE = 0.070, z = 0.73, p = .467, 95% CI [−0.086, 0.187]; n = 270 obs. (87 participants)
Enjoyment	Spearman	ρ = .523, p < .001, n = 289	b = 0.348, SE = 0.068, z = 5.14, p < .001, 95% CI [0.215, 0.481]; n = 198 obs. (105 participants)
B. Predictors of enjoyment			
Sex	Mann–Whitney U	U = 15,242.0, p = .776, r_{rb} = −.02; n = 352 (208 F, 144 M); males M = −0.019, SD = 0.961; females M = 0.012, SD = 1.033	b = 0.133, SE = 0.168, z = 0.79, p = .428, 95% CI [−0.196, 0.462]; n = 198 obs. (105 participants)
Age	Spearman	ρ = −.082, p = .126, n = 352	b = −0.006, SE = 0.007, z = −0.87, p = .384, 95% CI [−0.020, 0.008]; n = 198 obs. (105 participants)
AE	Spearman	ρ = .199, p < .001, n = 352; Holm-corrected p < .001	b = 0.179, SE = 0.065, z = 2.74, p = .006, 95% CI [0.051, 0.307]; n = 114 obs. (84 participants); Holm-corrected p = .012
PA	Spearman	ρ = .036, p = .501, n = 352	b = 0.082, SE = 0.082, z = 1.00, p = .320, 95% CI [−0.079, 0.243]; n = 114 obs. (84 participants)
MT	Spearman	ρ = .008, p = .888, n = 352	b = 0.076, SE = 0.046, z = 1.64, p = .101, 95% CI [−0.015, 0.166]; n = 114 obs. (84 participants)
SA	Spearman	ρ = −.022, p = .683, n = 352	b = 0.039, SE = 0.062, z = 0.63, p = .530, 95% CI [−0.083, 0.161]; n = 114 obs. (84 participants)
Em	Spearman	ρ = .113, p = .034, n = 352; Holm-corrected p = .034	b = 0.090, SE = 0.076, z = 1.19, p = .235, 95% CI [−0.058, 0.238]; n = 114 obs. (84 participants)
C. Multiple linear regression predicting retrospective PoT			
Enjoyment	OLS regression	b = 0.512, SE = 0.050, β = 0.52, t(286) = 10.28, p < .001, 95% CI [0.414, 0.610]	—
AE	OLS regression	b = 0.065, SE = 0.040, β = 0.08, t(286) = 1.64, p = .102, 95% CI [−0.013, 0.143]	—
Overall model	OLS regression	F(2, 286) = 60.76, p < .001, R^2 = .298, adjusted R^2 = .293; n = 289	—

Note. PoT and enjoyment were z-scored within experiment. Retrospective: Spearman correlations (ρ); two-sided Mann–Whitney U tests with rank-biserial correlation (r_{rb}) and M (SD) per sex; section C: ordinary least squares (OLS) regression, β = standardized coefficient. Prospective: linear mixed models (LMM) with a random intercept for participant, one predictor per model; b = fixed-effect estimate with standard error (SE), Wald z, two-sided p and 95% confidence interval (CI). Holm-corrected p: adjusted for the two Gold-MSI subscales tested in H3 and H4. Sex: female = reference (positive values = higher for males). n = participants; obs. = observations; AE = Active Engagement; PA = Perceptual Abilities; MT = Musical Training; SA = Singing Abilities; Em = Emotions; — = not calculated.

Table 5

Summary of Findings Regarding Duration Estimates and Accuracy

Predictor	Retrospective test	Retrospective result	Prospective result (LMM)
A. Predictors of duration estimates (bias towards under- or overestimation)			
Sex	Mann–Whitney U	U = 12,782.0, p = .016, r_{rb} = .15; n = 352 (206 F, 146 M); males M = 0.175, SD = 1.029; females M = −0.125, SD = 0.967	b = 0.250, SE = 0.103, z = 2.44, p = .015, 95% CI [0.049, 0.451]; n = 705 obs. (201 participants)
Age	Spearman	ρ = .143, p = .007, n = 352	b = −0.008, SE = 0.005, z = −1.59, p = .111, 95% CI [−0.018, 0.002]; n = 701 obs. (200 participants)
AE	Spearman	ρ = −.004, p = .942, n = 352	b = 0.024, SE = 0.042, z = 0.57, p = .570, 95% CI [−0.059, 0.106]; n = 621 obs. (180 participants)
PA	Spearman	ρ = .099, p = .063, n = 352	b = 0.072, SE = 0.047, z = 1.55, p = .122, 95% CI [−0.019, 0.163]; n = 621 obs. (180 participants)
MT	Spearman	ρ = .069, p = .199, n = 352	b = 0.092, SE = 0.028, z = 3.24, p = .001, 95% CI [0.037, 0.148]; n = 621 obs. (180 participants); Holm-corrected p = .002
SA	Spearman	ρ = .041, p = .447, n = 352	b = 0.047, SE = 0.041, z = 1.15, p = .248, 95% CI [−0.033, 0.127]; n = 621 obs. (180 participants)
Em	Spearman	ρ = .050, p = .349, n = 352	b = 0.010, SE = 0.039, z = 0.25, p = .801, 95% CI [−0.066, 0.086]; n = 621 obs. (180 participants)
Enjoyment	Spearman	ρ = −.130, p = .015, n = 348	b = −0.008, SE = 0.067, z = −0.12, p = .908, 95% CI [−0.138, 0.123]; n = 196 obs. (104 participants)
B. Predictors of duration-estimation accuracy (absolute error)			
Sex	Mann–Whitney U	U = 15,761.0, p = .442, r_{rb} = −.05; n = 352 (206 F, 146 M); males M = 0.232, SD = 0.182; females M = 0.233, SD = 0.181	b = −0.178, SE = 0.095, z = −1.87, p = .061, 95% CI [−0.365, 0.008]; n = 705 obs. (201 participants)
Age	Spearman	ρ = −.078, p = .144, n = 352	b = −0.002, SE = 0.005, z = −0.49, p = .623, 95% CI [−0.011, 0.007]; n = 701 obs. (200 participants)
AE	Spearman	ρ = −.015, p = .780, n = 352	b = 0.006, SE = 0.039, z = 0.14, p = .885, 95% CI [−0.071, 0.082]; n = 621 obs. (180 participants)
PA	Spearman	ρ = −.065, p = .224, n = 352	b = −0.059, SE = 0.043, z = −1.36, p = .172, 95% CI [−0.143, 0.026]; n = 621 obs. (180 participants)
MT	Spearman	ρ = −.007, p = .899, n = 352	b = −0.063, SE = 0.027, z = −2.38, p = .017, 95% CI [−0.116, −0.011]; n = 621 obs. (180 participants)
SA	Spearman	ρ = .038, p = .474, n = 352	b = −0.036, SE = 0.038, z = −0.96, p = .339, 95% CI [−0.110, 0.038]; n = 621 obs. (180 participants)
Em	Spearman	ρ = −.018, p = .736, n = 352	b = −0.020, SE = 0.036, z = −0.57, p = .569, 95% CI [−0.090, 0.049]; n = 621 obs. (180 participants)
Enjoyment	Spearman	ρ = .021, p = .702, n = 348	b = −0.015, SE = 0.067, z = −0.22, p = .822, 95% CI [−0.147, 0.117]; n = 196 obs. (104 participants)
C. Multiple linear regression predicting retrospective duration estimates			
Sex (male)	OLS regression	b = 0.330, SE = 0.107, β = 0.33, t(344) = 3.08, p = .002, 95% CI [0.120, 0.541]	—
Age	OLS regression	b = 0.007, SE = 0.004, β = 0.10, t(344) = 1.87, p = .063, 95% CI [−0.000, 0.014]	—
Enjoyment	OLS regression	b = −0.110, SE = 0.053, β = −0.11, t(344) = −2.07, p = .039, 95% CI [−0.215, −0.005]	—
Overall model	OLS regression	F(3, 344) = 5.92, p < .001, R^2 = .049, adjusted R^2 = .041; n = 348	—
D. Final linear mixed model predicting prospective duration estimates (sex + MT; ML estimation)			
Sex (male)	—	—	b = 0.261, SE = 0.106, z = 2.47, p = .014, 95% CI [0.053, 0.468]
MT	—	—	b = 0.093, SE = 0.028, z = 3.36, p < .001, 95% CI [0.039, 0.148]

Predictor	Retrospective test	Retrospective result	Prospective result (LMM)
Overall model	—	—	marginal R^2 = .046, conditional R^2 = .342; AIC = 1684.69 (MT only: 1688.67; sex only: 1693.62); n = 621 obs. (180 participants)

Note. Duration estimates = relative estimates (estimate/real duration), z-scored within experiment × paradigm; absolute error = |relative estimate − 1|, z-scored in the same way (lower = more accurate; means per sex in the Mann–Whitney tests are raw values). Tests, statistics, coding and abbreviations as in Table 4. Section D: LMM with both predictors (maximum likelihood); AIC = Akaike information criterion (lower = better fit), computed on the same observations; marginal and conditional R^2 = variance explained by the fixed effects and by the whole model. Holm-corrected p: adjusted for the two Gold-MSI subscales tested in H2.

Retrospective

Passage of time

A Mann-Whitney U test showed no significant effect for sex on passage of time judgements, U = 11531.5, p = .091, r_rb = -.11. A correlation matrix including PoT judgements, age, scores for the gold-MSI subscales, and enjoyment ratings, revealed significant effects on PoT judgements only for the subscale active engagement (AE), ρ = .20, p < .001, and enjoyment ratings, ρ = .52, p <.001. Results thus confirmed H7 and partially confirmed H3. Enjoyment was significantly correlated with scores for the AE and Em subscales (ρ = .20, p < .001, and ρ = .11, p = .0342, respectively).

To compare the strength of the effects of enjoyment and AE, a multiple linear regression was calculated having both of these as predictors for PoT judgements. According to results, enjoyment was the strongest predictor of retrospective PoT judgements, b = 0.512, SE = 0.050, t(286) = 10.28, p < .001, 95% CI [0.414, 0.610], with higher enjoyment being associated with a rating of a faster passage of time. In comparison, AE scores did not result in a significant effect, b = 0.065, SE = 0.040, t(286) = 1.64, p = .102, 95% CI [−0.013, 0.143], suggesting shared variance between these two factors. The overall regression model was significant, F(2, 286) = 60.76, p < .001, explaining 29.8% of the variance in perceived passage of time (R^2 = .298, adjusted R^2 = .293).

Duration estimates

Regarding retrospective DE, a Mann-Whitney test revealed a significant effect for sex U = 12782.0, p = .016, r_rb = .15. This contradicted H8. However, since the existing evidence suggests that sex-related effects on retrospective duration may be task-dependent, we ran the same test again, dividing our sample according to estimation method. This revealed that the sex difference was present for reproduction estimates, U = 3274.0, p = .001, r_rb = .28, but not for verbal estimates (Tempo Exp. 4), U = 3023.0, p = .962, r_rb = .00. Spearman correlations showed that retrospective DE were significantly and positively correlated with age, ρ = .14, p = .007, in the direction predicted by H5. Retrospective DE were significantly and negatively correlated with enjoyment ratings, ρ = -.13, p = .015, which was not previously hypothesized. None of the Gold-MSI subscales correlated

significantly with DE, contradicting H1. A regression analysis combining sex, enjoyment, and age as three possible predictors of standardized relative DE showed a significant effect for sex, $b = 0.330$, SE = 0.107, $t(344) = 3.08$, $p = .002$, 95% CI [0.120, 0.541] and enjoyment, $b = -0.110$, SE = 0.053, $t(344) = -2.07$, $p = .039$ 95% CI [−0.215, −0.005]. However, age was no longer significant, $b = 0.007$, SE = 0.004, $t(344) = 1.87$, $p = .063$. A Spearmam correlation confirmed age was not significantly correlated with enjoyment, ρ = -.082 p = .1255, and a Mann-Whitney test confirmed no significant age differences across sexes, U = 16488.0, $p = 0.2252$, r_rb = -.08. The coefficient for age was also the same when entered alone ($b = 0.007$, $p = .066$). Although the overall regression model was significant, $F(3, 344) = 5.92$, $p < .001$, it explained only 4.9% of the variance in relative DE ($R^2 = .049$, adjusted $R^2 = .041$).

In terms of accuracy, there was no significant effect of sex on absolute error, and no significant correlations were found between absolute error and scores for the Gold-MSI subscales, enjoyment or age.

Prospective

Passage of time

Neither age, sex nor any Gold-MSI subscale showed a statistically significant effect on PoT judgements (see Table 4). These results contradicted the effects hypothesized in H4, for the AE and Em subscales on PoT judgements. To further investigate the role of enjoyment stated in H4, a similar LMM was conducted, having enjoyment as the only fixed factor. It found a significant effect of enjoyment, $b = 0.348$, SE = 0.068, $z = 5.144$, $p < .001$, 95% CI [0.215, 0.481], with higher enjoyment being associated with the perception of time passing faster. Enjoyment alone accounted for 12% of the variance, and when combined with participants' intercept, for 34.76% (considering marginal and conditional R^2 values, respectively). Residuals 'distribution was normal, as confirmed by Shapiro-Wilk test (p = .847). This confirmed the effect of enjoyment hypothesized in H4. To further inspect the relationships between enjoyment, PoT judgements and musical expertise, five additional LMMs were fit, all having enjoyment as the dependent variable, participants as random factors, and each having one Gold-MSI subscale as a fixed factor. This revealed that AE scores were positively and significantly associated with higher enjoyment, $b = 0.179$, SE = 0.065, $z = 2.735$, $p = .006$, 95% CI [0.051, 0.307]. None of the remaining subscales showed a statistically significant effect on enjoyment.

In the models predicting prospective PoT judgements, participant random-intercept variance ranged from .235 to .453 across models, indicating substantial between-participant variation in passage-of-time responses.

Duration estimates

A procedure similar to the one conducted for PoT judgements was performed for prospective DE. This revealed that z-scored relative DE were significantly associated with sex, $b = 0.250$, SE = 0.103, $z = 2.44$, $p = .015$, 95% CI [0.049, 0.451]. This confirmed H9. Regarding musical expertise, a significant effect on prospective duration was found only for the MT subscale, $b = 0.092$, SE = 0.028, $z = 3.24$, $p = .001$, 95% CI [0.037, 0.148]. Higher musical training scores were associated with longer estimates, partially confirming H2. The remaining individual factors, also including enjoyment, did not impact DE significantly (see Table 5). This confirmed the null effect of age predicted in H6 but failed to confirm the effect of PA hypothesized in H2.

Sex and musical training were included as fixed factors together in a final model for DE. Both factors retained a significant effect: sex, $b = 0.261$, SE = 0.106, $z = 2.466$, $p = .014$, 95% CI [0.053, 0.468], and MT, $b = 0.093$, SE = 0.028, $z = 3.355$, $p < .001$, 95% CI [0.039, 0.148]. The values for marginal and conditional R^2 were .046 and .342, respectively, and AIC was lower with both factors included, 1684.692, than if including musical training (1688.674), or sex (1693.622) only.

To investigate the accuracy of prospective DE, similar models were fit, having absolute error as the dependent variable. This revealed a small but significant effect of the MT subscale, with higher training being associated with lower error, meaning higher accuracy, $b = -0.063$, SE = 0.027, $z = -2.382$, $p = .017$, 95% CI [−0.116, −0.011]. None of the remaining individual factors were significantly associated with the accuracy of prospective DE. Table 6 summarizes the final status of the initial hypotheses, against the reported findings.

Table 6
Summary of Hypotheses and Their Status

Hypothesis	Prediction	Status
H1	Higher scores for the PA and MT subscales will be negatively associated with DE, retrospectively.	Denied
H2	Higher scores for the PA and MT subscales will be positively associated with DE, prospectively.	Partially supported
H3	Higher scores for the Em and AE subscales will be associated with higher enjoyment of the music and with faster PoT, retrospectively.	Partially supported
H4	Higher scores for the Em and AE subscales will be associated with higher enjoyment of the music and with faster PoT, prospectively.	Partially supported
H5	Age will be positively associated with retrospective duration estimates.	Denied
H6	Age will not directly affect prospective duration estimates.	Confirmed
H7	Age will not directly affect retrospective judgements of the PoT.	Confirmed
H8	Females will retrospectively estimate duration as longer than males.	Denied
H9	Females will prospectively estimate duration as shorter than males.	Confirmed

Note. **Status according to the results reported in Tables 4 and 5. PA = Perceptual Abilities; MT = Musical Training; AE = Active Engagement; Em = Emotions; DE = duration estimates; PoT = passage of time.**

Discussion

In this study we investigated whether age, sex and musical expertise influenced time perception while listening to music. Findings confirmed two of our hypotheses related to age (H6 and H7) and to sex (H9) Partial support was found for three hypotheses related to musical expertise: H2, H3, and H4. Three hypotheses were denied by our findings: H1, H5, and H8.

Musical expertise

The hypothesis H1 predicted that participants with higher scores for the PA and MT subscales would provide shorter retrospective DE. This was not confirmed by the results, which found no significant effects on retrospective judgements. An additional analysis showed that higher scores of these subscales did not improve accuracy of DE. These findings contradict one previous piece of evidence [34], where musicians were found to provide shorter and more accurate retrospective DE than non-musicians. It could be argued that this difference was due to the method of collecting DE (only verbal in Phillips & Cross [34]; combination of verbal and reproduction in the present study). However, splitting our sample into retrospective verbal and reproduction estimates, and re-calculating correlations did not confirm this possibility. No significant correlations were found for any subscale when

filtering for either estimation method (Supplementary Table S1). An important aspect concerns how musical expertise was assessed. Our studies employed the Gold-MSI (mostly a reduced version). Phillips and Cross [34] measured expertise by years of formal training and instrumental practice. Future research should further investigate these findings, for instance, presenting the musical stimuli used in both studies, including both verbal and reproduction estimates for the same participants, and assessing musical expertise with the full Gold-MSI questionnaire.

In the prospective condition, higher scores for MT were associated with longer DE, partially supporting the H2. This suggests that participants with more extensive musical training may have been able to process music more efficiently, possibly freeing up attention to track time, leading to longer DE. Conversely, less experienced listeners may have needed to allocate more attention to processing the musical information, losing some mental 'pulses' (in the language of attentional gate models [7]) used to track time, leading to shorter estimates. Perceptual abilities were not found to influence prospective DE, contrary to what was hypothesized. Findings thus suggest that only the skills acquired through extended periods of systematic musical training may influence prospective DE, and not perceptual skills as self-evaluated by participants, and captured by the PA subscale. Importantly, the effects found for MT reflected both a shift in estimation bias and an improvement in accuracy. Since most prospective estimates were shorter than the real duration (64% of trials), longer estimates tend to be closer to the real duration, so the accuracy effect may follow from the shift in bias rather than reflect a separate improvement in timing precision. A positive relationship between accuracy and musical training was previously found in one study [35], but contradicted by another [20]. Further research is needed to understand whether music sophistication affects prospective duration estimation accuracy, and under which situations those effects might occur.

In the case of the perceived PoT, the results confirmed the prominent role of enjoyment hypothesized in H3 and H4. These findings align with existing evidence from empirical research using stimuli beyond music listening [13]. Our hypotheses also predicted that higher scores in the AE and Em subscales would lead to greater enjoyment and correlate with PoT judgements, which was not fully confirmed. While higher scores of AE were associated with higher enjoyment, both retrospectively and prospectively, only in the retrospective paradigm was AE positively correlated with PoT. Em was significantly associated with enjoyment only retrospectively, and with neither paradigm was it directly associated with PoT judgements. One explanation could be that certain emotions experienced during music listening might have a biographical dimension, and occur in response to episodic memories evoked by music [57]. These particular individual emotional responses to music would be captured by the Em subscale, but would likely be more prominent when listening to certain familiar musical pieces, as opposed to the newly-composed stimuli used in these experiments. Participants with higher Em scores in our sample might have experienced greater distortion in perceived PoT if familiar music had been used. In combination, our data points to the possibility that participants with higher AE scores, who

engage more actively and frequently with music, tend to enjoy the music more, which then would lead them to perceive time as passing faster.

Age

Even though age was positively correlated with retrospective duration (H5), when it was included in a multiple linear regression together with sex and enjoyment, it no longer exerted a significant effect on retrospective DE. Its coefficient did not change when sex and enjoyment were added, so the different outcomes of the Spearman correlation and the regression reflect the different tests rather than an overlap with these factors. A similar effect could have happened in Phillips and Cross' [34] experiment (whose findings informed H5), had their analyses considered age as a continuous variable (instead of age categories) and included a combination of age, enjoyment and sex in a comprehensive predictive model. Further research is needed to confirm whether age is a relevant factor for retrospective estimates of musical duration, especially considering the likely effects of sex and enjoyment.

Regarding prospective estimates, our results confirm H6, suggesting no effects of age. These findings add to existing evidence reporting no age-related differences in prospective DE [38-40]. Regarding PoT judgements, results confirmed H7, showing no effects of age. This replicates previous evidence employing non-musical tasks with a short duration range [38, 45].

Sex

Our results showed a significant effect of sex on DE, for retrospective and prospective paradigms, with females providing shorter DE than males (retrospectively, only for reproduction estimates). These findings contradict our hypothesis for retrospective duration (H8), but support our hypothesis in relation to prospective duration (H9).

In the case of the retrospective condition, the sex difference appeared only for reproduction estimates (r_rb = .28) and was absent for verbal estimates (r_rb = .00; Supplementary Table S1). Females' longer retrospective estimates in the meta-analysis by Block et al. [46] came from verbal estimation, the method for which we found no sex difference. Together with the prospective results, this suggests that males' longer estimates could be linked to the reproduction task itself rather than to how much musical information was memorized. Because estimation method, stimulus duration and sample were confounded in our retrospective data, this interpretation remains tentative. It remains possible that females indeed memorized less musical information than males, or that music perception required less cognitive effort for them, and that such difference may have been masked due to the quantization of responses that tends to happen with the verbal method, and may be more significant for longer time intervals [38,58,59]. Since participants were not asked about the

amount of information memorized or the cognitive effort experienced, these possible explanations cannot be investigated. Future research should further investigate whether sex may influence retrospective DE differently for music listening than for other tasks, and whether such an effect interacts with the task, or the duration of the stimulus. In the retrospective paradigm sex-related differences appear to be restricted to a bias in temporal estimation, with no evidence of an effect in relation to accuracy of DE.

In the prospective condition, results supported H9, with females providing shorter durations than males. According to the existing literature, these findings might suggest either a different attentional behaviour, with females paying less attention to time, and/or a faster internal clock for females, restricted to the reproduction task. It is important to note that an overall faster baseline rate of the internal clock for females could not explain our findings under a reproduction method. Females' internal clocks would simply accumulate more pulses during both the listening and the reproduction stage, compared to males. An explanation based on internal clock differences would only hold in the case that the females' internal clocks run faster only in the reproduction step, leading them to reproduce the perceived duration as shorter. This is still possible and could reflect a sex-dependent arousal effect associated with the performance of the reproduction task. Future research could investigate sex-related effects on time perception, combining the collection of bodily measures (e.g., looking at galvanic skin response, which can give a measure of physiological arousal) with subjective ratings of attention, for different types of tasks. Our exploratory findings regarding PoT revealed no effects of sex in either paradigm, paving the way for future confirmatory research in this field.

Enjoyment and duration

One exploratory finding of the current study suggests that enjoyment was associated with shorter retrospective DE. Even though this was a small effect, it is a notable relationship, considering that in previous evidence, enjoyment is mostly associated with PoT judgements, and not DE. It is possible that lower enjoyment led to the experience of boredom, with participants naturally focusing on duration more, even though the task itself was retrospective. This could have led them to provide longer DE than participants who enjoyed the music and only after listening to it estimated duration retrospectively. This explanation cannot be tested with our data. Future experiments should employ a retrospective timing task and ask participants about enjoyment and how much attention they felt that they dedicated to tracking time.

Individual baseline differences

The mixed-effects models indicated substantial between-participant variability in baseline temporal judgements, beyond the variance accounted for by the three main individual factors considered in this study. Future research should attempt to directly measure other sources of variance in time perception already identified in the literature, such as level of attention, memory and cognitive load, and evaluate their impacts in a music listening context. Besides these variables, other individual factors have been suggested as possibly impacting time perception, such as personality type and cultural background. For instance, more impulsive individuals may experience boredom more easily, possibly leading to a slower perceived speed of the PoT, and longer DE [60]. Higher interoceptive awareness of visceral sensations, such as heart rate and respiration, increases the accuracy of prospective reproductions of duration [61] especially in the presence of irregular auditory stimuli [62]. Handedness has also been suggested as a possible influence on time perception [47].

Limitations

This study pooled data from seven experiments. Although stimuli and procedures were considered comparable and all measures were standardized within experiment, the final sample was not collected as part of a single study. First, in the retrospective paradigm the differences in estimation method were combined with a difference in the range of stimulus duration and instrumentation: all verbal estimates came from one experiment (Tempo Exp. 4) with 3.5–4.5-min versions of a piano piece, while all reproduction estimates relate to 16–38-s melodies played by a flute. Differences between methods, such as the sex difference found only for reproduction, therefore cannot be attributed to any one of these factors. Second, musical sophistication was assessed with the complete Gold-MSI in one experiment and with reduced versions in the others, which may have lowered the reliability of subscale scores. Third, part of the sample contributed to both paradigms in the same session, so retrospective and prospective results are not fully independent (this was taken account of in the analysis). Finally, the individual factors explained little variance in DE (marginal $R^2 \leq .05$), whereas stable differences between participants indicated that other individual characteristics may play a larger role.

Conclusion

The analysis of a combined dataset pooled from seven experiments on music and time perception suggested that individual differences play a modest role in how listeners judge musical time. Enjoyment was the strongest influence on the felt passage of time in both paradigms (prospective and retrospective), and scores for self-reported Active Engagement with musical activities were associated with greater enjoyment. Males gave longer DE than females in both paradigms, although retrospectively, this was found only for reproduction, not for verbal estimates. The amount of musical training was positively associated with

longer and more accurate prospective estimates. Overall, age was shown to not be a significant factor for prospective or retrospective judgements of musical duration or passage of time. Individual factors explained less variance than stable baseline differences between participants, suggesting that other individual characteristics may play a stronger role in time perception in musical contexts, and should be targets for future research.

References

1. Zakay, D. & Block, R. A. Prospective and retrospective duration judgments: an executive-control perspective. *Acta Neurobiol. Exp.* **64**, 319–328 (2004).
2. Zakay, D. & Block, R. A. Temporal cognition. *Curr. Dir. Psychol. Sci.* **6**, 12–16 (1997).
3. Wittmann, M. The inner experience of time. *Phil. Trans. R. Soc. B* **364**, 1955–1967 (2009).
4. Ornstein, R. E. *On the Experience of Time* (Penguin Books, 1969).
5. Poynter, D. Judging the duration of time intervals: a process of remembering segments of experience in *Time and Human Cognition: A Life-Span Perspective* (eds Levin, I. & Zakay, D.) 305–331 (Elsevier, 1989). https://doi.org/10.1016/S0166-4115(08)61045-6
6. Gibbon, J., Church, R. M. & Meck, W. Scalar timing in memory. *Ann. N. Y. Acad. Sci.* **423**, 52–77 (1984).
7. Zakay, D. & Block, R. A. An attentional-gate model of prospective time estimation in *Time and the Dynamic Control of Behavior* (eds Richelle, M., De Keyser, V., d'Ydewalle, G. & Vandierendonck, A.) 167–178 (Université de Liège, 1995).
8. Jones, L. A. The perception of duration and the judgment of the passage of time in *The Illusions of Time: Philosophical and Psychological Essays on Timing and Time Perception* (eds Arstila, V., Bardon, A., Power, S. E. & Vatakis, A.) (Palgrave Macmillan, 2019). https://doi.org/10.1007/978-3-030-22048-8_4
9. Penton-Voak, I. S., Edwards, H., Percival, A. & Wearden, J. Speeding up an internal clock in humans? Effects of click trains on subjective duration. *J. Exp. Psychol. Anim. Behav. Process.* **22**, 307–320 (1996). https://doi.org/10.1037/0097-7403.22.3.307
10. Zakay, D., Nitzan, D. & Glicksohn, J. The influence of task difficulty and external tempo on subjective time estimation. *Percept. Psychophys.* **34**, 451–456 (1983). https://doi.org/10.3758/BF03203060
11. Droit-Volet, S. & Wearden, J. H. Experience sampling methodology reveals similarities in the experience of passage of time in young and elderly adults. *Acta Psychol.* **156**, 77–82 (2015). https://doi.org/10.1016/j.actpsy.2015.01.006
12. Droit-Volet, S. & Wearden, J. Passage of time judgments are not duration judgments: evidence from a study using experience sampling methodology. *Front. Psychol.* **7**, 176 (2016). https://doi.org/10.3389/fpsyg.2016.00176
13. Martinelli, N. & Droit-Volet, S. Judgment of duration and passage of time in prospective and retrospective conditions and its predictors for short and long durations. *Sci. Rep.* **12**, 22241 (2022). https://doi.org/10.1038/s41598-022-25913-9
14. Droit-Volet, S., Ramos, D., Bueno, J. L. O. & Bigand, E. Music, emotion, and time perception: the influence of subjective emotional valence and arousal? *Front. Psychol.* **4**, 417 (2013). https://doi.org/10.3389/fpsyg.2013.00417
15. Silva, L. B., Phillips, M. & Martins, J. O. The influence of tonality, tempo, and musical sophistication on the listener's time-duration estimates. *Q. J. Exp. Psychol.* **77**, 1846–1864 (2024). https://doi.org/10.1177/17470218231203459
16. Bailey, N. & Areni, C. S. Background music as a quasi clock in retrospective duration judgments. *Percept. Mot. Skills* **102**, 435–444 (2006).

17. Phillips, M. et al. What determines the perception of segmentation in contemporary music? *Front. Psychol.* **11**, 1001 (2020). https://doi.org/10.3389/fpsyg.2020.01001
18. Silva, L. B., Bellmann, O. T., Phillips, M. & Martins, J. O. Musical timbre influences time perception: differential effects of human and artificial voices. *Music Percept.* 1–21 (2026). https://doi.org/10.1525/mp.2026.2618953
19. Bailey, N. & Areni, C. S. When a few minutes sound like a lifetime: does atmospheric music expand or contract perceived time? *J. Retail.* **82**, 189–202 (2006). https://doi.org/10.1016/j.jretai.2006.05.003
20. Hammerschmidt, D. & Wöllner, C. Sensorimotor synchronization with higher metrical levels in music shortens perceived time. *Music Percept.* **37**, 263–277 (2020). https://doi.org/10.1525/MP.2020.37.4.263
21. Pereira, L. A. S., Ramos, D. & Bueno, J. L. O. The influence of different musical modes and tempi on time perception. *Acta Psychol.* **229**, 103701 (2022). https://doi.org/10.1016/j.actpsy.2022.103701
22. Ziv, N. & Omer, E. Music and time: the effect of experimental paradigm, musical structure and subjective evaluations on time estimation. *Psychol. Music* **39**, 182–195 (2010). https://doi.org/10.1177/0305735610372612
23. Jones, M. R. Musical time in *The Oxford Handbook of Music Psychology* (eds Hallam, S., Cross, I. & Thaut, M.) 125–142 (Oxford University Press, 2016).
24. Jones, L. A. An overview of the psychology of time perception in *Music and Time: Psychology, Philosophy, Practice* (eds Phillips, M. & Sergeant, M.) 42–54 (Boydell & Brewer, 2022).
25. Müllensiefen, D., Gingras, B., Musil, J. & Stewart, L. The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PLoS ONE* **9**, e89642 (2014). https://doi.org/10.1371/journal.pone.0089642
26. Levitin, D. J., Grahn, J. A. & London, J. The psychology of music: rhythm and movement. *Annu. Rev. Psychol.* **69**, 51–75 (2018). https://doi.org/10.1146/annurev-psych-122216-011740
27. Repp, B. H. Sensorimotor synchronization and perception of timing: effects of music training and task experience. *Hum. Mov. Sci.* **29**, 200–213 (2010). https://doi.org/10.1016/j.humov.2009.08.002
28. Scheurich, R., Zamm, A. & Palmer, C. Tapping into rate flexibility: musical training facilitates synchronization around spontaneous production rates. *Front. Psychol.* **9**, 458 (2018). https://doi.org/10.3389/fpsyg.2018.00458
29. Grondin, S. *The Perception of Time: Your Questions Answered* (Routledge, 2020).
30. Wang, X., Burger, B. & Wöllner, C. Tapping to drumbeats in an online experiment changes our perception of time and expressiveness. *Psychol. Res.* **88**, 127–140 (2024). https://doi.org/10.1007/s00426-023-01835-7
31. Panagiotidi, M. & Samartzi, S. Time estimation: musical training and emotional content of stimuli. *Psychol. Music* **41**, 620–629 (2013). https://doi.org/10.1177/0305735612441737
32. Dowling, W. J., Bartlett, J. C., Halpern, A. R. & Andrews, M. W. Melody recognition at fast and slow tempos: effects of age, experience, and familiarity. *Percept. Psychophys.* **70**, 496–502 (2008). https://doi.org/10.3758/PP.70.3.496
33. Talamini, F., Altoè, G., Carretti, B. & Grassi, M. Musicians have better memory than nonmusicians: a meta-analysis. *PLoS ONE* **12**, e0186773 (2017). https://doi.org/10.1371/journal.pone.0186773
34. Phillips, M. & Cross, I. About musical time – effect of age, enjoyment, and practical musical experience on retrospective estimate of elapsed duration during music listening in *Multidisciplinary Aspects of Time and Time Perception* (eds Vatakis, A., Esposito, A., Giagkou, M., Cummins, F. & Papadelis, G.) 125–136 (Springer, 2011). https://doi.org/10.1007/978-3-642-21478-3_11
35. Hammerschmidt, D., Wöllner, C., London, J. & Burger, B. Disco time: the relationship between perceived duration and tempo in music. *Music Sci.* **4**, 1–11 (2021). https://doi.org/10.1177/2059204320986384
36. Ericsson, A. & Pool, R. *Peak: Secrets from the New Science of Expertise* (Houghton Mifflin Harcourt, 2016).
37. Liao, H.-W., Hommelhoff, S. & Carstensen, L. L. Time perception from seconds to lifetimes: how perceived time affects adult development in *The Cambridge Handbook of Cognitive Aging: A Life Course Perspective* (eds Thomas, A. K. & Gutchess, A.) 254–272 (Cambridge University Press, 2020).
38. Wearden, J. *The Psychology of Time Perception* (Palgrave Macmillan, 2016). https://doi.org/10.1057/978-1-137-40883-9

39. Lamotte, M. & Droit-Volet, S. Aging and time perception for short and long durations: a question of attention? *Timing Time Percept.* **5**, 149–167 (2017). https://doi.org/10.1163/22134468-00002086
40. Ragot, R., Ferrandez, A.-M. & Pouthas, V. Time, music, and aging. *Psychomusicology* **18**, 28–45 (2002). https://doi.org/10.1037/h0094053
41. Szelag, E. & Skolimowska, J. Cognitive function in elderly can be ameliorated by training in temporal information processing. *Restor. Neurol. Neurosci.* **30**, 419–434 (2012). https://doi.org/10.3233/RNN-2012-120240
42. Hammerschmidt, D. Spontaneous motor tempo in *Performing Time: Synchrony and Temporal Flow in Music and Dance* (eds Wöllner, C. & London, J.) 141–150 (Oxford University Press, 2023).
43. El Haj, M., Moroni, C., Samson, S., Fasotti, L. & Allain, P. Prospective and retrospective time perception are related to mental time travel: evidence from Alzheimer's disease. *Brain Cogn.* **83**, 45–51 (2013). https://doi.org/10.1016/j.bandc.2013.06.008
44. Friedman, W. J. & Janssen, S. M. J. Aging and the speed of time. *Acta Psychol.* **134**, 130–141 (2010). https://doi.org/10.1016/j.actpsy.2010.01.004
45. Wittmann, M. & Lehnhoff, S. Age effects in perception of time. *Psychol. Rep.* **97**, 921–935 (2005). https://doi.org/10.2466/pr0.97.3.921-935
46. Block, R. A., Hancock, P. A. & Zakay, D. Sex differences in duration judgments: a meta-analytic review. *Mem. Cognit.* **28**, 1333–1346 (2000). https://doi.org/10.3758/BF03211834
47. Koglbauer, I. Gender differences in time perception in *The Cambridge Handbook of Applied Perception Research* (eds Hoffman, R. R., Hancock, P. A., Scerbo, M. W., Parasuraman, R. & Szalma, J. L.) 1004–1028 (Cambridge University Press, 2015). https://doi.org/10.1017/CBO9780511973017.059
48. Espinosa-Fernández, L., Miró, E., Cano, M. & Buela-Casal, G. Age-related changes and gender differences in time estimation. *Acta Psychol.* **112**, 221–232 (2003). https://doi.org/10.1016/S0001-6918(02)00093-8
49. Glicksohn, J. & Hadad, Y. Sex differences in time production revisited. *J. Individ. Differ.* **33**, 35–42 (2012). https://doi.org/10.1027/1614-0001/a000059
50. Hancock, P. A. & Rausch, R. The effects of sex, age, and interval duration on the perception of time. *Acta Psychol.* **133**, 170–179 (2010). https://doi.org/10.1016/j.actpsy.2009.11.005
51. Hanson, N. J. & Buckworth, J. Sex differences in time perception during self-paced running. *Int. J. Exerc. Sci.* **9**, 514–523 (2016). https://doi.org/10.70252/XRJV6155
52. Castillo-Escamilla, J., Salvador-Viñas, M. d. M. & Cimadevilla, J. M. Sex based differences in time processing were found in newly developed tasks. *Psychol. Res.* **90**, 44 (2026). https://doi.org/10.1007/s00426-026-02244-2
53. Giovannelli, F. et al. Gender differences in time perception during olfactory stimulation. *J. Sens. Stud.* **31**, 61–69 (2016). https://doi.org/10.1111/joss.12191
54. Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
55. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).
56. Schielzeth, H. et al. Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods Ecol. Evol.* **11**, 1141–1152 (2020). https://doi.org/10.1111/2041-210X.13434
57. Juslin, P. N. *Musical Emotions Explained: Unlocking the Secrets of Musical Affect* (Oxford University Press, 2019).
58. Block, R. A., Grondin, S. & Zakay, D. Prospective and retrospective timing processes: theories, methods, and findings in *Timing and Time Perception: Procedures, Measures, & Applications* (eds Vatakis, A., Balci, F., di Luca, M. & Correa, Á.) 32–51 (Brill, 2018). https://doi.org/10.1163/9789004280205_003
59. Ogden, R. S., Simmons, F. R. & Wearden, J. H. Verbal estimation of the magnitude of time, number, and length. *Psychol. Res.* **85**, 3048–3060 (2021). https://doi.org/10.1007/s00426-020-01456-4
60. Wittmann, M., Jokic, T. & Pfeifer, E. Modulations in the experience of duration in *The Illusions of Time: Philosophical and Psychological Essays on Timing and Time Perception* (eds Arstila, V., Bardon, A., Power, S. E. & Vatakis, A.) 145–162 (Palgrave Macmillan, 2019).

61. Meissner, K. & Wittmann, M. Body signals, cardiac awareness, and the perception of time. *Biol. Psychol.* **86**, 289–297 (2011). https://doi.org/10.1016/j.biopsycho.2011.01.001

62. Teghil, A., Boccia, M., Nocera, L., Pietranelli, V. & Guariglia, C. Interoceptive awareness selectively predicts timing accuracy in irregular contexts. *Behav. Brain Res.* **377**, 112242 (2020). https://doi.org/10.1016/j.bbr.2019.112242

Author contributions

LS, JW, MP and LJ jointly conceptualized and planned one or more of the experiments. LS and JW worked on data collection and preparation. LS analyzed the data and drafted the manuscript. JW revised the manuscript. LS and MP acquired funding.

Data availability statement

The full dataset and analysis script are available at https://doi.org/10.6084/m9.figshare.34856070. The musical stimuli presented in the experiments are available upon request.

Ethics Statements

The experiments included in this study received ethical approval by the Centre for Interdisciplinary Studies, based in University of Coimbra, Portugal. We followed the European Charter of Fundamental Rights, worked along the guidelines of "Good Scientific Practice," and fulfilled the ethical principles for research involving human subjects (Helsinki declaration). Informed consent was obtained online from all participants, prior to the beginning of each experiment.

Funding

This study received funding from FCT - Fundação para a Ciencia e a Tecnologia, I.P. through the Centre for Interdisciplinary Studies, in the scope of the project UID/00460/2025 - https://doi.org/10.54499/UID/ 00460/2025, and by the PhD grant 2023.01403.BD, with DOI https://doi.org/10.54499/2023.01403.BD, attributed to L.B.S. Additional funding to compensate participants was also provided by the Royal Northern College of Music.

Confict of Interests

The authors declare no competing interests.

Tables

Table 1

Number of Participants and Observations per Experiment of Origin

Experiment of origin	Participants	Retro. DE	Retro. PoT	Prosp. DE	Prosp. PoT	Gold-MSI
Tempo Exp. 3	132	128	131	—	—	Reduced
Tonality Exp. 2 [15]	116	63	—	353	—	Reduced
Tempo Exp. 4 (verbal estimates)	161	161	161	—	—	Complete (38 items)
Timbre Exp. 1 [18]	73	—	—	73	73	Reduced
Timbre Exp. 2 [18]	21	—	—	84	84	Not collected
Tonality Exp. 3	41	—	—	39	41	Reduced
Meter	56	—	—	156	156	Reduced
Total	437 (unique)	352	292	705	354	

Note. Counts after removal of outlying trials and of the 2 participants who reported another sex category or did not answer. Prospective cells count observations (trials); several participants contributed up to four trials and took part in more than one experiment. DE = duration estimates; PoT = passage of time; — = not collected. Reduced = Gold-MSI versions with two to four items per subscale; Complete = full 38-item inventory.

Table 2

Demographic Characteristics and Gold-MSI Subscale Scores of Participants

Variable	*n*	*M*	*SD*	Min	25th	Mdn	75th	Max
Age (years)	436	36.36	14.42	18.00	24.00	32.00	44.00	78.00
Active Engagement	416	4.15	1.27	1.00	3.33	4.00	5.00	7.00
Perceptual Abilities	416	4.94	1.07	1.00	4.33	5.00	5.67	7.00
Musical Training	416	2.82	1.71	1.00	1.29	2.29	4.00	7.00
Singing Abilities	416	4.01	1.36	1.00	3.00	4.00	5.00	7.00
Emotions	416	5.42	1.21	1.00	4.50	5.50	6.50	7.00
Sex								
Female	251	57.4%						
Male	186	42.6%						

Note. $N = 437$ participants included in the analyses. Gold-MSI subscale scores (range 1–7) were not collected in Timbre Exp. 2, hence the smaller n. M = mean; SD = standard deviation; Min and Max = observed range; 25th, Mdn and 75th = percentiles. For Sex, the M column shows the percentage of participants.

Table 3
Musical Characteristics of Presented Stimuli

Experiment of origin	Instrumentation	Name of Stimulus	Tempo (BPM)	Duration
Tempo Exp. 3	Flute	Melody A	60	35 s
			150	38 s
Tonality Exp. 2 [15]	Flute	Melody A	120	19 s
			90	26 s
		Melody B	120	16 s
			90	21 s
Tempo Exp. 4	Piano	Time to Go (by Matthew Woolhouse)	86	267 s
			98	236 s
			110	211 s
Timbre Exp. 1 [18]	Female voice	Natural melody	90	31 s
Timbre Exp. 2 [18]	Female voice	Happy 1	70	38 s
		Happy 2	100	31 s
		Sad 1	66	39 s
		Sad 2	100	32 s
Tonality Exp. 3	Flute	Melody B - slow	90	21 s
Meter	Female voice + piano	Melody with bass in quarter-notes	120	20 s
		Melody with bass in eighth-notes	120	20 s
		Melody with bass in half-notes	120	20 s

Note. BPM = beats per minute; s = seconds.

Table 4

Summary of Findings Regarding Passage of Time and Enjoyment

Predictor	Retrospective test	Retrospective result	Prospective result (LMM)
A. Predictors of passage of time (PoT) judgements			
Sex	Mann–Whitney U	U = 11,531.5, p = .091, r_{rb} = −.11; n = 292 (171 F, 121 M); males M = −0.104, SD = 0.969; females M = 0.073, SD = 1.021	b = −0.230, SE = 0.146, z = −1.58, p = .115, 95% CI [−0.517, 0.056]; n = 354 obs. (108 participants)
Age	Spearman	ρ = −.017, p = .769, n = 292	b = 0.000, SE = 0.007, z = 0.04, p = .965, 95% CI [−0.013, 0.014]; n = 354 obs. (108 participants)
AE	Spearman	ρ = .197, p < .001, n = 292; Holm-corrected p = .001	b = 0.030, SE = 0.063, z = 0.48, p = .630, 95% CI [−0.094, 0.154]; n = 270 obs. (87 participants)
PA	Spearman	ρ = .001, p = .984, n = 292	b = 0.045, SE = 0.075, z = 0.59, p = .553, 95% CI [−0.103, 0.192]; n = 270 obs. (87 participants)
MT	Spearman	ρ = .014, p = .815, n = 292	b = 0.050, SE = 0.044, z = 1.13, p = .260, 95% CI [−0.037, 0.136]; n = 270 obs. (87 participants)
SA	Spearman	ρ = −.034, p = .567, n = 292	b = −0.072, SE = 0.059, z = −1.22, p = .221, 95% CI [−0.186, 0.043]; n = 270 obs. (87 participants)
Em	Spearman	ρ = .040, p = .494, n = 292	b = 0.051, SE = 0.070, z = 0.73, p = .467, 95% CI [−0.086, 0.187]; n = 270 obs. (87 participants)
Enjoyment	Spearman	ρ = .523, p < .001, n = 289	b = 0.348, SE = 0.068, z = 5.14, p < .001, 95% CI [0.215, 0.481]; n = 198 obs. (105 participants)
B. Predictors of enjoyment			
Sex	Mann–Whitney U	U = 15,242.0, p = .776, r_{rb} = −.02; n = 352 (208 F, 144 M); males M = −0.019, SD = 0.961; females M = 0.012, SD = 1.033	b = 0.133, SE = 0.168, z = 0.79, p = .428, 95% CI [−0.196, 0.462]; n = 198 obs. (105 participants)
Age	Spearman	ρ = −.082, p = .126, n = 352	b = −0.006, SE = 0.007, z = −0.87, p = .384, 95% CI [−0.020, 0.008]; n = 198 obs. (105 participants)
AE	Spearman	ρ = .199, p < .001, n = 352; Holm-corrected p < .001	b = 0.179, SE = 0.065, z = 2.74, p = .006, 95% CI [0.051, 0.307]; n = 114 obs. (84 participants); Holm-corrected p = .012
PA	Spearman	ρ = .036, p = .501, n = 352	b = 0.082, SE = 0.082, z = 1.00, p = .320, 95% CI [−0.079, 0.243]; n = 114 obs. (84 participants)
MT	Spearman	ρ = .008, p = .888, n = 352	b = 0.076, SE = 0.046, z = 1.64, p = .101, 95% CI [−0.015, 0.166]; n = 114 obs. (84 participants)
SA	Spearman	ρ = −.022, p = .683, n = 352	b = 0.039, SE = 0.062, z = 0.63, p = .530, 95% CI [−0.083, 0.161]; n = 114 obs. (84 participants)
Em	Spearman	ρ = .113, p = .034, n = 352; Holm-corrected p = .034	b = 0.090, SE = 0.076, z = 1.19, p = .235, 95% CI [−0.058, 0.238]; n = 114 obs. (84 participants)
C. Multiple linear regression predicting retrospective PoT			
Enjoyment	OLS regression	b = 0.512, SE = 0.050, β = 0.52, t(286) = 10.28, p < .001, 95% CI [0.414, 0.610]	—
AE	OLS regression	b = 0.065, SE = 0.040, β = 0.08, t(286) = 1.64, p = .102, 95% CI [−0.013, 0.143]	—
Overall model	OLS regression	F(2, 286) = 60.76, p < .001, R^2 = .298, adjusted R^2 = .293; n = 289	—

Note. PoT and enjoyment were z-scored within experiment. Retrospective: Spearman correlations (ρ); two-sided Mann–Whitney U tests with rank-biserial correlation (r_{rb}) and M (SD) per sex; section C: ordinary least squares (OLS) regression, β = standardized coefficient. Prospective: linear mixed models (LMM) with a random intercept for participant, one predictor per model; b = fixed-effect estimate with standard error (SE), Wald z, two-sided p and 95% confidence interval (CI). Holm-corrected p: adjusted for the two Gold-MSI subscales tested in H3 and H4. Sex: female = reference (positive values = higher for males). n = participants; obs. = observations; AE = Active Engagement; PA = Perceptual Abilities; MT = Musical Training; SA = Singing Abilities; Em = Emotions; — = not calculated.

Table 5

Summary of Findings Regarding Duration Estimates and Accuracy

Predictor	Retrospective test	Retrospective result	Prospective result (LMM)
A. Predictors of duration estimates (bias towards under- or overestimation)			
Sex	Mann–Whitney U	U = 12,782.0, p = .016, r_{rb} = .15; n = 352 (206 F, 146 M); males M = 0.175, SD = 1.029; females M = −0.125, SD = 0.967	b = 0.250, SE = 0.103, z = 2.44, p = .015, 95% CI [0.049, 0.451]; n = 705 obs. (201 participants)
Age	Spearman	ρ = .143, p = .007, n = 352	b = −0.008, SE = 0.005, z = −1.59, p = .111, 95% CI [−0.018, 0.002]; n = 701 obs. (200 participants)
AE	Spearman	ρ = −.004, p = .942, n = 352	b = 0.024, SE = 0.042, z = 0.57, p = .570, 95% CI [−0.059, 0.106]; n = 621 obs. (180 participants)
PA	Spearman	ρ = .099, p = .063, n = 352	b = 0.072, SE = 0.047, z = 1.55, p = .122, 95% CI [−0.019, 0.163]; n = 621 obs. (180 participants)
MT	Spearman	ρ = .069, p = .199, n = 352	b = 0.092, SE = 0.028, z = 3.24, p = .001, 95% CI [0.037, 0.148]; n = 621 obs. (180 participants); Holm-corrected p = .002
SA	Spearman	ρ = .041, p = .447, n = 352	b = 0.047, SE = 0.041, z = 1.15, p = .248, 95% CI [−0.033, 0.127]; n = 621 obs. (180 participants)
Em	Spearman	ρ = .050, p = .349, n = 352	b = 0.010, SE = 0.039, z = 0.25, p = .801, 95% CI [−0.066, 0.086]; n = 621 obs. (180 participants)
Enjoyment	Spearman	ρ = −.130, p = .015, n = 348	b = −0.008, SE = 0.067, z = −0.12, p = .908, 95% CI [−0.138, 0.123]; n = 196 obs. (104 participants)
B. Predictors of duration-estimation accuracy (absolute error)			
Sex	Mann–Whitney U	U = 15,761.0, p = .442, r_{rb} = −.05; n = 352 (206 F, 146 M); males M = 0.232, SD = 0.182; females M = 0.233, SD = 0.181	b = −0.178, SE = 0.095, z = −1.87, p = .061, 95% CI [−0.365, 0.008]; n = 705 obs. (201 participants)
Age	Spearman	ρ = −.078, p = .144, n = 352	b = −0.002, SE = 0.005, z = −0.49, p = .623, 95% CI [−0.011, 0.007]; n = 701 obs. (200 participants)
AE	Spearman	ρ = −.015, p = .780, n = 352	b = 0.006, SE = 0.039, z = 0.14, p = .885, 95% CI [−0.071, 0.082]; n = 621 obs. (180 participants)
PA	Spearman	ρ = −.065, p = .224, n = 352	b = −0.059, SE = 0.043, z = −1.36, p = .172, 95% CI [−0.143, 0.026]; n = 621 obs. (180 participants)
MT	Spearman	ρ = −.007, p = .899, n = 352	b = −0.063, SE = 0.027, z = −2.38, p = .017, 95% CI [−0.116, −0.011]; n = 621 obs. (180 participants)
SA	Spearman	ρ = .038, p = .474, n = 352	b = −0.036, SE = 0.038, z = −0.96, p = .339, 95% CI [−0.110, 0.038]; n = 621 obs. (180 participants)
Em	Spearman	ρ = −.018, p = .736, n = 352	b = −0.020, SE = 0.036, z = −0.57, p = .569, 95% CI [−0.090, 0.049]; n = 621 obs. (180 participants)
Enjoyment	Spearman	ρ = .021, p = .702, n = 348	b = −0.015, SE = 0.067, z = −0.22, p = .822, 95% CI [−0.147, 0.117]; n = 196 obs. (104 participants)
C. Multiple linear regression predicting retrospective duration estimates			
Sex (male)	OLS regression	b = 0.330, SE = 0.107, β = 0.33, t(344) = 3.08, p = .002, 95% CI [0.120, 0.541]	—
Age	OLS regression	b = 0.007, SE = 0.004, β = 0.10, t(344) = 1.87, p = .063, 95% CI [−0.000, 0.014]	—
Enjoyment	OLS regression	b = −0.110, SE = 0.053, β = −0.11, t(344) = −2.07, p = .039, 95% CI [−0.215, −0.005]	—
Overall model	OLS regression	F(3, 344) = 5.92, p < .001, R^2 = .049, adjusted R^2 = .041; n = 348	—
D. Final linear mixed model predicting prospective duration estimates (sex + MT; ML estimation)			
Sex (male)	—	—	b = 0.261, SE = 0.106, z = 2.47, p = .014, 95% CI [0.053, 0.468]
MT	—	—	b = 0.093, SE = 0.028, z = 3.36, p < .001, 95% CI [0.039, 0.148]

Predictor	Retrospective test	Retrospective result	Prospective result (LMM)
Overall model	—	—	marginal R^2 = .046, conditional R^2 = .342; AIC = 1684.69 (MT only: 1688.67; sex only: 1693.62); n = 621 obs. (180 participants)

Note. Duration estimates = relative estimates (estimate/real duration), z-scored within experiment × paradigm; absolute error = |relative estimate − 1|, z-scored in the same way (lower = more accurate; means per sex in the Mann–Whitney tests are raw values). Tests, statistics, coding and abbreviations as in Table 4. Section D: LMM with both predictors (maximum likelihood); AIC = Akaike information criterion (lower = better fit), computed on the same observations; marginal and conditional R^2 = variance explained by the fixed effects and by the whole model. Holm-corrected p: adjusted for the two Gold-MSI subscales tested in H2.

Table 6

Summary of Hypotheses and Their Status

Hypothesis	Prediction	Status
H1	Higher scores for the PA and MT subscales will be negatively associated with DE, retrospectively.	Denied
H2	Higher scores for the PA and MT subscales will be positively associated with DE, prospectively.	Partially supported
H3	Higher scores for the Em and AE subscales will be associated with higher enjoyment of the music and with faster PoT, retrospectively.	Partially supported
H4	Higher scores for the Em and AE subscales will be associated with higher enjoyment of the music and with faster PoT, prospectively.	Partially supported
H5	Age will be positively associated with retrospective duration estimates.	Denied
H6	Age will not directly affect prospective duration estimates.	Confirmed
H7	Age will not directly affect retrospective judgements of the PoT.	Confirmed
H8	Females will retrospectively estimate duration as longer than males.	Denied
H9	Females will prospectively estimate duration as shorter than males.	Confirmed

Note. Status according to the results reported in Tables 4 and 5. PA = Perceptual Abilities; MT = Musical Training; AE = Active Engagement; Em = Emotions; DE = duration estimates; PoT = passage of time.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1SpearmanCorrelationsRetrospectiveDuration.docx