UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia Generale

Corso di laurea Magistrale in Psicologia Cognitiva Applicata

Tesi di laurea Magistrale

Exploring the relationship between music consonance and tonal fusion

Relatore
Prof. M. Grassi

Laureanda: Maria Giulia Levorin
Matricola: 2050626

Anno Accademico 2022/2023

Table of Contents

1 - Theoretical Introduction

In this chapter, some key concepts – tonal consonance, tonal dissonance, and tonal fusion – are introduced and compared within the main theoretical proposals.

1.1 - Musical Consonance or Psychological Dissonance?

Playing two music tones simultaneously, or a more complex harmonic interval, two principles are definable in Western tonal music: tonal consonance, which describes the level of stability and pleasantness experienced by the listeners, and tonal dissonance, that defines the level of instability and unpleasantness perceived (Yuan, 2020). The concept of musical consonance stems from ancient Greek philosophy. It was believed that certain mathematical ratios between tones induced consonance – pleasurableness – whilst other ratios created dissonance in harmonic intervals, as they were unpleasurable (Terhardt, 1984).

Since the time of Pythagoras, different accounts for tonal consonance or dissonance have been proposed (DeWitt & Crowder, 1987). The notation system and the theory of harmony developed during the Middle Ages were based on consonance and dissonance principles. Music theorists such as Guido of Arezzo[1] and Boethius[2] thought that tonal consonance was typical of tones arranged in perfect intervals – fourth, fifth and octave – whilst imperfect intervals – thirds and sixths – produced tonal dissonance (Bowling & Purves, 2015). Renaissance theorists believed that major seconds, minor seconds and tritone were the musical intervals related to dissonance, while major thirds, minor thirds and the so-considered *perfect intervals* were related to consonance (Di Stefano

---

[1] Musician and theorist Guido of Arezzo (born in Arezzo circa 992, died: 1050). He created a musical pedagogy based on rational experience and practical training of the singer. He invented solmisation and improved staff notation (Enciclopedia Treccani, 2023a) Anicius Manlius Severinus Boethius (born: circa 475–7 C.E., died: 526? C.E.) was one of the most important intermediaries between ancient Greek philosophy and the Middle Ages; he made philosophical ideas accessible to a wider public. All his writings drew on Greek Neoplatonists (John Marenbon, 2021)..

[2] Anicius Manlius Severinus Boethius (born: circa 475–7 C.E., died: 526? C.E.) was one of the most important intermediaries between ancient Greek philosophy and the Middle Ages; he made philosophical ideas accessible to a wider public. All his writings drew on Greek Neoplatonists (John Marenbon, 2021).

& Bertolaso, 2014). Different modern disciplines were interested in musical consonance – for instance physics, physiology, mathematics, and psychology – as the full understanding of this perceptual phenomenon is still a challenge (Di Stefano & Bertolaso, 2014).

An important point of view was offered by Hermann von Helmholtz[3]. He proposed a sensory theory of consonance, focused on hearing processes and physiology. Consonance perception in concurrent tones would depend on the simplicity of mathematical ratios between tones' frequencies. He compared ear functioning with a piano, considering a set of fibers – located in the inner ear – that responded to the various frequencies (Martinelli, 2014).

Then, Carl Stumpf[4] promoted an alternative explanation – more focused on perception than the one taken by Helmholtz – to the basis of tonal consonance. He proposed that various grades of tonal fusion in an harmonic interval could predict tonal consonance; more tonal fusion would predict more tonal consonance (DeWitt & Crowder, 1987).

1.1.1 – Like Kant's Blue Lenses: An Ethnomusicology Perspective

Consonance or dissonance are used in music across cultures with different purposes (Cook, 2021). In Western classical music, consonance is apt to create a sense of resolution or stability, whilst dissonance conveys a sense of unease or tension. Also, Indian classical music uses consonance to convey harmonicity and stability. In some African music traditions, consonance is linked to the sense of community – unity – leading to rhythmic cohesion. In those traditions, dissonance is implemented to create a sense of energy and excitement. Chinese traditional music considers consonance to convey a sense of joy or happiness. Chinese stringed instruments *guzheng* and *guqin*

---

[3] Hermann von Helmholtz wrote the book (1863/1954) "On the Sensations of Tone as a Physiological Basis for the Theory of Music" that was the major reference for musical perception for decades, and he conducted somehow controlled experiments (Yost, 2015).

[4] Carl Stumpf (1848-1936) was an anti-naturalist philosopher interested in tonal perception. In the two volumes of *Tonpsychologie* (*Psychology of Sound*) – written in 1883 and 1890 respectively – he proposed his idea of tonal fusion (*Tonverschmelzung*) and its role in consonance perception (Martinelli, 2014).

are often associated with consonant and dissonant intervals respectively; the latter is used to convey a sense of melancholy and sadness.

So, there are many cultural differences in the perception of musical consonance and dissonance. First of all, while in Western music the dichotomy between consonance and dissonance is emphasized, in other cultures it might not be so relevant (Wood & Lindsay, 1957). The perception of consonance or dissonance itself could change between cultures, due to musical training and cultural expectations. For instance, a study compared Balinese *gamelan* music with Western music in consonance and dissonance perception (Wendt & Bader, 2019). Gamelan music is an Indonesian traditional music style that uses a variety of percussion instruments with a tone system different from the Western "well-tempered" system. In this study, Balinese listeners generally rated the music as less dissonant than Western listeners.

Therefore, the concepts of musical consonance and dissonance in many cultures have expressive purposes; their perception and applications vary cross-culturally.

1.1.2 – Psychological Standpoint

From a psychological point of view, musical consonance seems related to the formation and violation of expectations in music (Sears, 2008). As a melody unfolds, listeners form expectations about incoming events. If these expectations are violated, then – for instance – a sense of tension could be created. This process can be experimentally tested with harmonic priming paradigms, where listeners are primed with a chord, thus creating expectations for the incoming next chord. If the expected chord is not presented, then the violation could convey a sense of tension, or non-resolution. Consonance and dissonance experiences might be related to the development – through passive music exposure – of implicit knowledge about musical expectations. However, the relationship between these variables is still not clear (Sears, 2008).

Considering just simultaneous consonance – perceived pleasantness of simultaneously sounding musical tones (Harrison & Pearce, 2019) – as a composite phenomenon, it seems to be related to cultural familiarity, interference and periodicity/harmonicity (Harrison & Pearce, 2019).

1.2 - Tonal Fusion

1.2.1 – What is tonal fusion?

Tonal fusion is a perceptual phenomenon that can be defined as "the mixing of at least two tones into a sole tonal experience" (Sam M.S., 2013). Over time, there have been various operationalizations of the concept[5]. Towards the end of the nineteenth century, Carl Stumpf proposed his influential definition of tonal fusion. He was inspired by Aristotelian tradition, as well as the perceptual experiments on sensations mixing by Ernst H. Weber[6]. His definition of tonal fusion (a) is the following:

(a) "Fusion is that *relationship* of two sense-perceived elements in which they form not a mere sum but rather a *whole*[7]. The result of this relationship is that with greater levels of fusion the overall impression, under otherwise identical circumstances, approaches closer and closer to the one of a *single perception*, and becomes increasingly difficult to break down." (Martinelli, 2014)

In his experiments, Stumpf registered how often tone pairs were misinterpreted as single tones. He proposed tonal fusion as fundamental in consonance perception (Martinelli, 2014). However, he also suggested that consonance and fusion could both be broadly linked to harmonicity recognition (Schneider, 1997). His contribution was fundamental in starting a research line on tonal fusion, musical consonance, and their relationship with harmonicity perception.

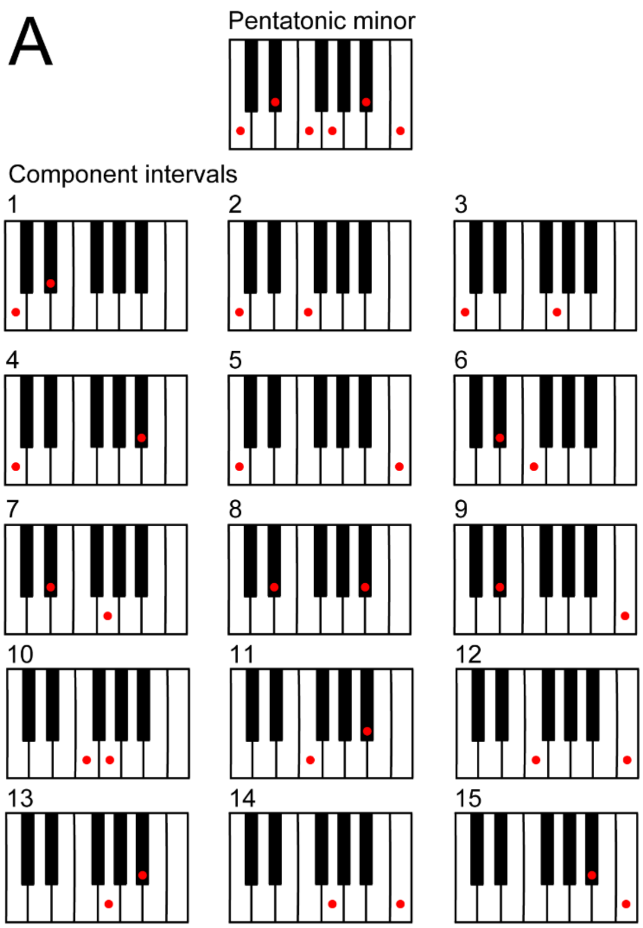
1.2.2 – When does tonal fusion happen?

Assuming the definition proposed by Stumpf (a), tonal fusion can be found using different experimental methods; with pure tones, but also with concurrent complex

---

[5] E.g., tonal fusion has been considered as *timbre* or *sonance* (Metfessel, 1926).
[6] Ernst H. Weber (1795-1878) physiologist, anatomist and psychologist, one of the fathers of modern psychophysics (Enciclopedia Treccani, 2023b).
[7] This concept of *whole* goes towards *Gestalt* theory (Martinelli, 2014).

tones with complementary partials (Huron, 1991). For instance, one of the first experiments examining its relation with consonance (DeWitt & Crowder, 1987) studied tonal fusion on three-tone combinations, when they created a perfect fifth or an octave interval in particular. Octave (1:2) is regarded as the most "fused" interval after the unison (1:1) prototypically. Tonal fusion has also been observed in more complex musical contexts, as Bach's polyphonic compositions (Huron, 1991).



B

Interval	Name	Ratio
1	Minor third	6:5
2	Perfect fourth	4:3
3	Perfect fifth	3:2
4	Minor seventh	9:5
5	Octave	2:1
6	Major second	10:9
7	Major third	5:4
8	Perfect fifth	3:2
9	Major sixth	5:3
10	Major second	9:8
11	Perfect fourth	4:3
12	Perfect fifth	3:2
13	Minor third	6:5
14	Perfect fourth	4:3
15	Major second	9:8

Fig.1: (A) Representation of music intervals on a pentatonic minor scale, (B) some intervals' names and their correspondent ratios. (Gill & Purves, 2009). Western music is based on a relatively small number of intervals, some of them are: unison (1:1), minor second (18:17), major second (9:8), minor third (6:5), major third (5:4), perfect fourth (4:3), tritone (7:5), perfect fifth (3:2), minor sixth (27:17), major sixth (5:3), minor seventh (9:5), major seventh (15:8), and perfect eighth or octave (2:1) (Dawson, 2010).

1.3 – Tonal fusion vs. Tonal consonance

Mixed evidence describes the relationship between tonal fusion and tonal consonance. As noticed by Harrison & Pearce (2020), in experiments with single tones and a requested judgement on the level of perceived fusion (DeWitt & Crowder, 1987; Guernsey, 1928) traditionally consonant intervals in western music – such as octaves and perfect fifths – were more often misperceived as a single tone than other intervals. This evidence is coherent the first Stumpf idea of a possible causation of consonance by tonal fusion. In polyphonic compositions, the number of vertical intervals resulted directly correlated to tonal consonance, except when they displayed tonal fusion (Huron, 1991). This dissociation between tonal fusion and consonance is coherent with the need of polyphonic music to maintain distinguishable contrapuntal voices. Furthermore, when tonal fusion was tested on chords with a pitch-matching procedure (McLachlan et al., 2013) the accuracy in matching a probe tone to a specified chord tone increased for consonant chords. This would suggest that the relationship between tonal fusion and tonal consonance is inverse to the one originally proposed by Stumpf (Harrison & Pearce, 2019).

1.3.1 – Experimental approach

Two experiments have approached tonal fusion and tonal consonance directly.

DeWitt & Crowder (1987) investigated tonal fusion as it was conceived by Stumpf. They operationalized tonal fusion as an increase in reaction times and/or errors to conventionally consonant intervals in a discrimination task. Their study comprehended three different experiments. In the first experiment, participants listened to single tones or tone pairs representing twelve tone intervals, and they had to indicate if they heard one or two tones. In the second experiment, participants were presented with a smaller set of intervals, and they had to distinguish between two-tones and three-tones combinations. The procedure was the same for the third experiment, where just note intervals in stimuli changed, to account for a possible explanation of fusion in terms of harmonic series. Overall, results suggest that the harmonic series is central to tonal fusion, with octave as the most fused interval, followed by fifth and fourth. They broadly support a relationship between tonal fusion and tonal consonance. However they do not exclude influences from other external factors in consonance perception, for instance prior expectations (DeWitt & Crowder, 1987).

McPherson et al. (2020) conducted cross-cultural experiments. Tonal fusion was considered together with aesthetic responses, in a native Amazonian society in Bolivia, the *Tsimane'* as well as in a sample of Westerners from Boston. Tsimane' tribe has been chosen because they have a limited exposure to Western music, and because group music performances are not in their culture. They do not have great experience with concurrent music pitches, and they seem to lack a preference for consonant chords (McDermott et al., 2016). In the proposed fusion experiments, participants had to judge whether the stimuli were composed of one or two sounds. These stimuli were composed of concurrent notes in different intervals, both from synthetic tones and sung notes. There was also a pleasantness rating task to account for consonance judgements, along with other control tasks designed to monitor Tsimanes' comprehension. It was found that both Westerners and Amazonians were more likely to consider as fused canonically consonant intervals, however only Boston's participants shown preference for those intervals. The fusion patterns seem consistent with harmonic series. Octave interval had the greatest fusion rate, followed by the fifth, then the fourth, and then the third.

1.4 – Actual experiment idea

The debate on the possible role of tonal fusion in consonance perception has not found a landing point yet. Therefore, considering the measures of tonal fusion and consonance in literature, an experiment to further investigate the relationship between those phenomena is proposed in this dissertation. In particular, the two studies described above (DeWitt & Crowder, 1987; McPherson et al., 2020) have been taken into account to design the present experiment. The first part measure reaction times and response accuracy (DeWitt & Crowder, 1987), whilst the second and third parts consider a consonance judgement (McPherson et al., 2020) and an original fusion judgement on a four points Likert scale. This aims to investigate tonal fusion perception not just as a dichotomy (e.g., one or two sounds), but also take in account a diverse judgement on what participants deem tonal fusion.

2 – Method

2.1 – Sample

Forty-four participants were recruited through snowball sampling. They gave their informed consent to enter the experiment. Their age was between 19 and 60 years old (mean 27.68 yr., SD 11.24 yr.). A total of twenty-three participants identified as females, seventeen as males, three as non-binary/third gender, one preferred not to say. Two participants were not born in Italy, having Romanian and Albanian origin respectively. All declared to have spent most of their life and childhood in Italy. Most of the participants declared to have high education grades; three had a master's degree, twenty had a three-years degree or professional qualification, twelve had upper secondary education, three had middle school qualifications, and six declared to be still in the educative system.

Participants compiled the Goldsmith Musical Sophistication Index (Gold-MSI) Questionnaire. The index ranges from 0 to 144 points, and it provides a quantitative measure of an individual's musical sophistication (Müllensiefen et al., 2013). Participants' general score was distributed from 21 to 100 points (60.89 mean score, 17.04 SD)[8]. Figure 2 shows score distributions for the main index, in Appendix A there are the representations of the five subscales distributions, while in Table 1 other main features are represented.

	MIN	MAX	MEDIAN	MEAN	SD
MSI INDEX	21	100	59.5	60.89	17.04
ACTIVE ENGAGEMENT	14	45	30	30.43	7.4
PERCEPTUAL ABILITIES	23	61	44.5	43.23	9.62
MUSICAL TRAINING	2	14	6.5	6.27	3.4
EMOTIONS	16	41	34	33.23	5.11
SINGING ABILITIES	8	46	30	28.41	9.03

Table 1: Gold-MSI Index Score main index and subscales.

[8] Average population score in the original validation study is 81.58 (Müllensiefen et al., 2013).

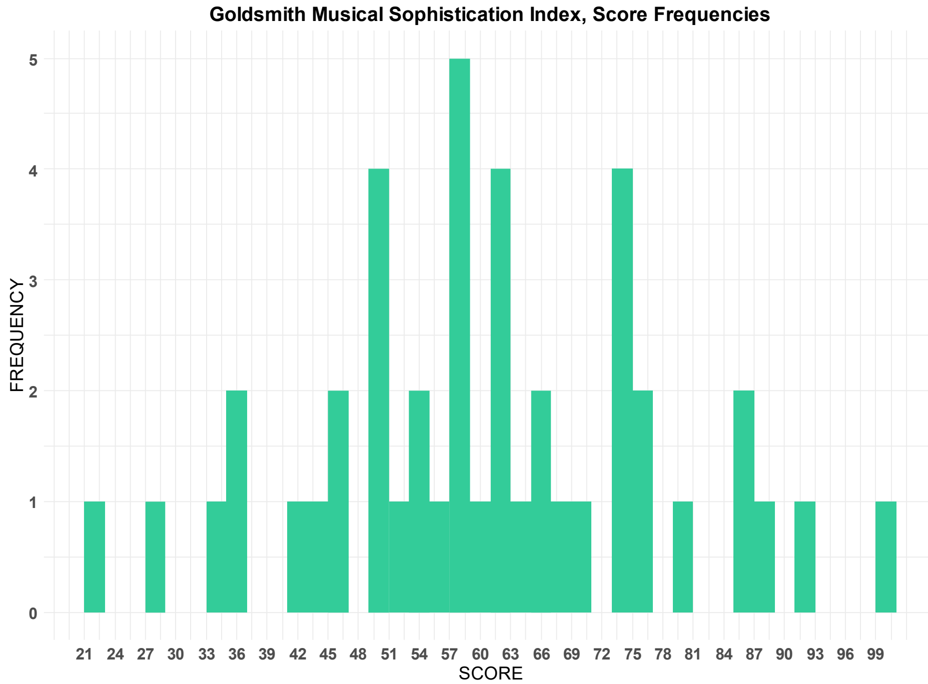
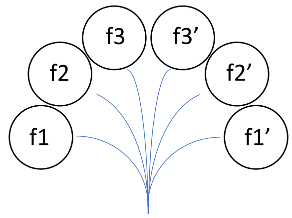


Fig.2: Gold-MSI Index Score frequency, represented by gender.

2.2 – Stimuli

Audio stimuli were created using MLP toolbox in MATLAB® considering synthetic tones in just intonation already used in literature (McPherson et al., 2020). Chosen three base frequencies – 110 Hz, 220 Hz and 440 Hz – the frequencies of two semitones above each of them were also selected, thus creating six base frequencies. From every frequency, nine different pitch intervals (Fig. 3) were created using two notes. Each synthetic tone had a duration of 2000 ms. Imitating real music instruments, a harmonic amplitudes attenuation by −14 dB/octave and a decaying exponential ($4s^{-1}$ decay constant) were applied. A temporal envelope raised cosine gates of 10 ms each at the onset and offset of the sound.



INTERVALS

1:1	9:8	5:4	4:3	45:32	3:2	15:18	1:2	32:15	**RATIO**
UNISON	MAJOR SECOND	MAJOR THIRD	PERFECT FOURTH	TRITONE	PERFECT FIFTH	MAJOR SEVENTH	OCTAVE	MAJOR NINTH	
0	203	386	498	590	702	1088	1200	1312	**CENTS**

Fig.3: note intervals created for every base frequency, with their ratios, also expressed in cents (just intonation).

2.3 – Procedure

The experiment was composed by three main tasks, followed by the Goldsmith Musical Sophistication Index Questionnaire[9] (Gold-MSI). Written instructions were given and explained to the participants. Each trial presented a sound interval, preceded by a fixation cross that disappeared with the audio onset. The participant listened to the audio via headphones and the response was collected using the keyboard[10]. The required response varied depending on the task.

First task asked the participant to answer rapidly if they had perceived one or two sounds. Half of the participants in the first half of the task had to press “z” key for one sound, “m” key for two sounds, whilst in the second part “z” key stood for two sounds and “m” key for one sound; vice versa for the remaining participants, counterbalancing trials.

[9] The Gold-MSI was administered using Qualtrics (Qualtrics, 2020). The Italian translation has not been validated yet. The translation applied in this study is retrievable in Appendix B.
[10] Details on hardware and software equipment, as well as MATLAB (The MathWorks Inc., 2023) are available in Appendices A, B, C.

Second and third task were also counterbalanced in their order of presentation. Participants answered with keys from "1" to "4". One task was a fusion judgement. They had to judge how much the perceived sound was fused, from "1" (not at all) to "4" (completely). The other task was a consonance judgement; it was asked how much they liked the sound, from "1" (dislike a lot) to "4" (like a lot). The order of task presentation thus differed across participants, that were divided in four different groups as shown in Table 2.

GROUP	TASK 1 - 1st part	TASK 1 - 2nd part	TASK 2	TASK 3
A	"z" one sound "m" two sounds	"m" one sound "z" two sounds	Fusion judgement	Consonance judgement
B	"z" one sound "m" two sounds	"m" one sound "z" two sounds	Consonance judgement	Fusion judgement
C	"m" one sound "z" two sounds	"z" one sound "m" two sounds	Fusion judgement	Consonance judgement
D	"m" one sound "z" two sounds	"z" one sound "m" two sounds	Consonance judgement	Fusion judgement

Table 2: groups of participants distinguished by tasks' presentation order.

First task was composed by 108 trials each, while second and third task had both 48 trials, due to the combination of base frequencies and interval range applied in the sound stimuli. The unison interval was used just in the first task – a control condition – as the participants had to distinguish between one and two sounds. This judgement is a fusion measure introduced in a previous study (McPherson et al., 2020).

In the middle of each task was provided to the participant the possibility to take a self-paced pause, as well as in both the parts of the first task. At the end of the tasks, participants compiled the Gold-MSI. The whole experiment required about 45 min to be completed.

Fig. 4 shows the general structure of a single trial.

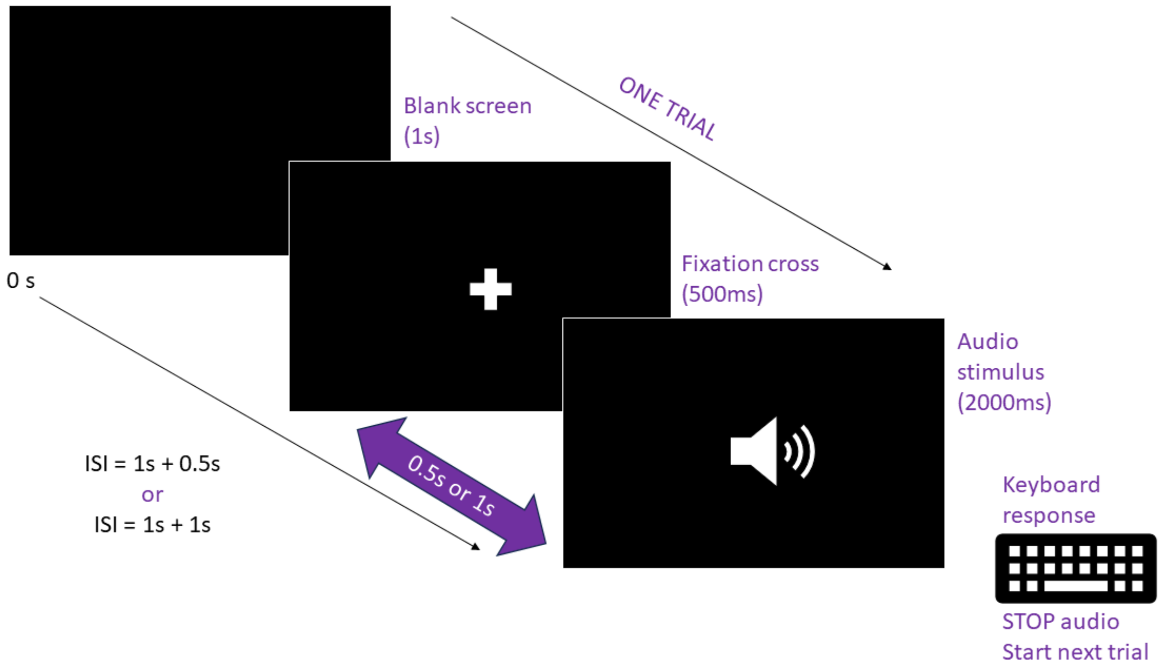


Fig. 4: overview of a single trial structure. The trial started with a blank screen that lasted 1s. Then the fixation cross appeared in the middle of the screen and lasted 0.5s. When the fixation cross disappeared, there was a brief pause without sound that had variable duration (0.5s or 1s), before the onset of the audio stimulus. This was planned to diminish anticipation and decreasing of attention in the participant. So, the Inter Stimulus Interval varied randomly between 1.5s and 2s, comprehending both the time (1s) before the fixation cross and before the audio onset (0.5s or 1s). Each audio stimulus had a duration of 2s, and it was repeated ten times for every task. When the participant pressed a key, the audio – if it was still going on – was interrupted and the next trial started. In the first task, if the participant did not press any key after 3s from the sound onset, the next trial started automatically.

3 – Analysis

3.1 – Descriptive statistics

Considering data from the first task, mean values of number of errors – where the participant answered one sound whilst an interval was presented, or vice versa – and reaction times were calculated for every interval (Table 3a, Fig. 5-6a). Trials where the response time was greater than 3s were excluded from the analysis, as well as trial where the participant did not give their response.

	ERRORS		REACTION TIMES	
INTERVAL IN CENTS	MEAN	SD	MEAN in ms	SD
0-203	0.2	0.4	1595.443	330.73
0-386	0.31	0.46	1614.653	341.98
0-498	0.34	0.48	1624.529	331.43
0-590	0.26	0.44	1603.335	322.49
0-702	0.5	0.5	1620.658	329.61
0-1088	0.23	0.42	1597.747	329.37
0-1200	0.7	0.45	1627.145	334.91
0-1312	0.22	0.42	1587.665	331.32
REACTION TIMES	1^{st} 1363.3 Q 3^{rd} 1836.9 Q	Median 1622.9	Mean 1604.0	132.1 min 2500.3 Max

Table 3a: means and standard deviations for errors in response accuracy and reaction times in the first task, all groups. Octave (highlighted in purple) had the larger mean number of errors and reaction times, followed in the error mean by the perfect fifth (highlighted in pink). The lowest mean numbers of errors and reaction times (highlighted in green and blue respectively) were for major second and major ninth. First and third quantiles, median, mean and range of total reaction times are in the lowest part of the table.

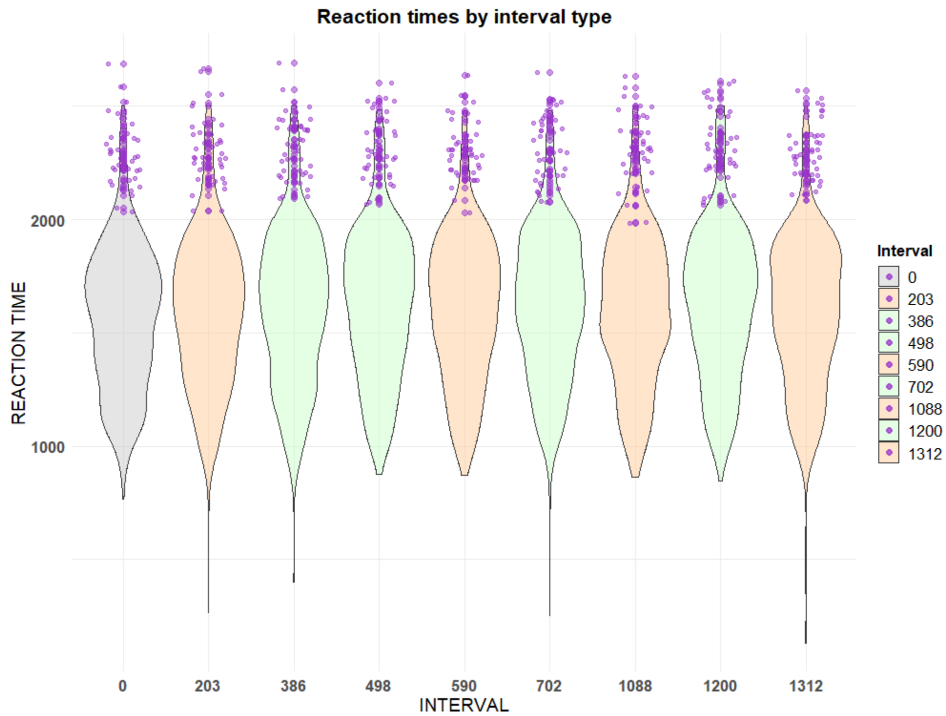


Fig. 5: Reaction times by interval type in the first task, all groups.

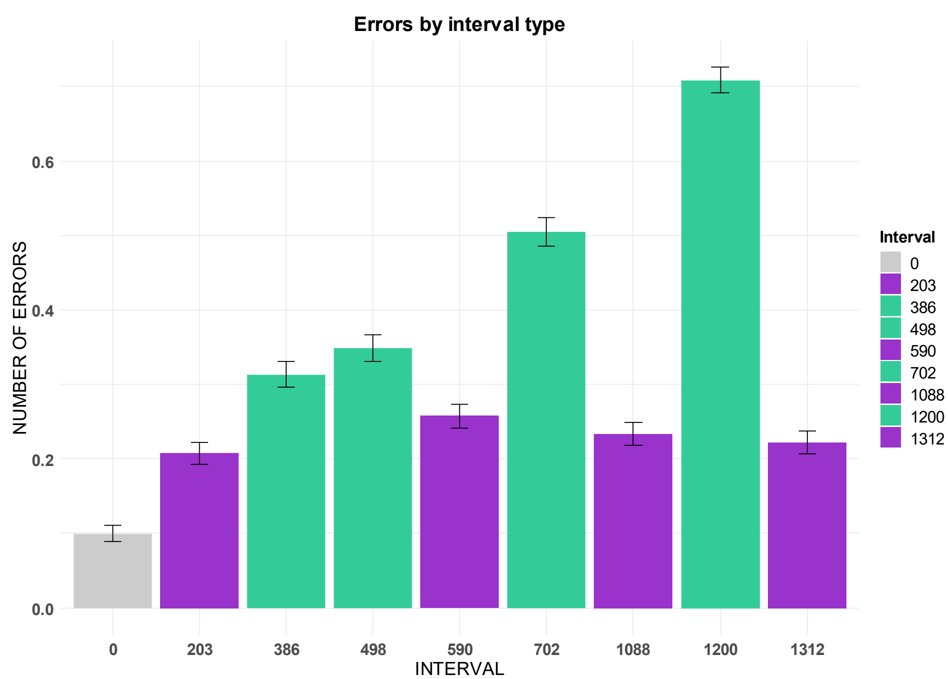


Fig. 6a: Errors by interval type in the first task considering all the sample. Dissonant intervals are represented in violet, consonant intervals in green and unison in grey.

Mean and standard deviation for both error and reaction times were observed also in contingence of the Gold-MSI global score. So, participants were subsampled in two groups, considering lower (Low Gold-MSI group) or higher (High Gold-MSI group) scores than the median in the Index (59.5 median). The values for these subgroups are represented in Table 3b-c and Figures 6b-c-d-e. The average number of errors was the highest for octave interval in both groups Low Gold-MSI and High Gold-MSI. Octave was followed by the perfect fifth interval in mean number of errors. The interval with less errors and smaller reaction times was major second for Low Gold-MSI group, and major ninth for High Gold-MSI group, thus differing from the overall sample.

L Subsample	ERRORS		REACTION TIMES	
INTERVAL IN CENTS	MEAN	SD	MEAN in ms	SD
0-203	0.19	0.39	1630.321	330.93
0-386	0.37	0.48	1676.864	338.49
0-498	0.43	0.50	1671.632	323.04
0-590	0.29	0.45	1678.715	300.01
0-702	0.58	0.49	1669.293	326.27
0-1088	0.28	0.45	1660.416	312.97
0-1200	0.74	0.44	1656.698	333.20
0-1312	0.27	0.44	1635.178	326.70
REACTION TIMES	1^{st} 1423.2 Q 3^{rd} 1874.2 Q	Median 1664.3	Mean 1654.7	252.6 min 2500.1 Max

Table 3b: Low Gold-MSI group means and standard deviations calculated for errors in response accuracy and reaction times in the first task. Octave (highlighted in purple) had the larger mean number of errors followed by perfect fifth (in pink), but the longest mean reaction times were for tritone (yellow). The

lowest mean number of errors and reaction times (in green) was for major second interval. First and third quantiles, median, mean and range of total reaction times are in the lowest part of the table.

H Subsample	ERRORS		REACTION TIMES	
INTERVAL IN CENTS	MEAN	SD	MEAN in ms	SD
0-203	0.22	0.41	1568.202	328.40
0-386	0.27	0.44	1567.496	337.45
0-498	0.28	0.45	1587.801	333.64
0-590	0.23	0.42	1549.543	327.51
0-702	0.44	0.50	1581.951	327.53
0-1088	0.20	0.40	1551.250	333.86
0-1200	0.68	0.47	1601.944	334.74
0-1312	0.18	0.39	1549.207	330.42
REACTION TIMES	1st 1308.0 Q 3rd 1808.2 Q	Median 1578.8	Mean 1564.5	132.1 min 2500.3 Max

Table 3c: High Gold-MSI group mean and standard deviations calculated for errors in response accuracy and reaction times in the first task. Octave (highlighted in purple) had the larger mean number of errors and mean reaction times, followed by perfect fifth in error means (highlighted in pink) and perfect fourth (in yellow) for reaction times. The lowest mean number of errors and reaction times (highlighted in green) was for major ninth interval. First and third quantiles, median, mean and range of total reaction times are in the lowest part of the table.

Reaction times range differed in the two groups. While in Low Gold-MSI group the minimum value was 252.6 ms, in High Gold-MSI group the minimum was 132.1 ms. Also mean, median and the first quantile are lower in High Gold-MSI group than in Low Gold-MSI group.

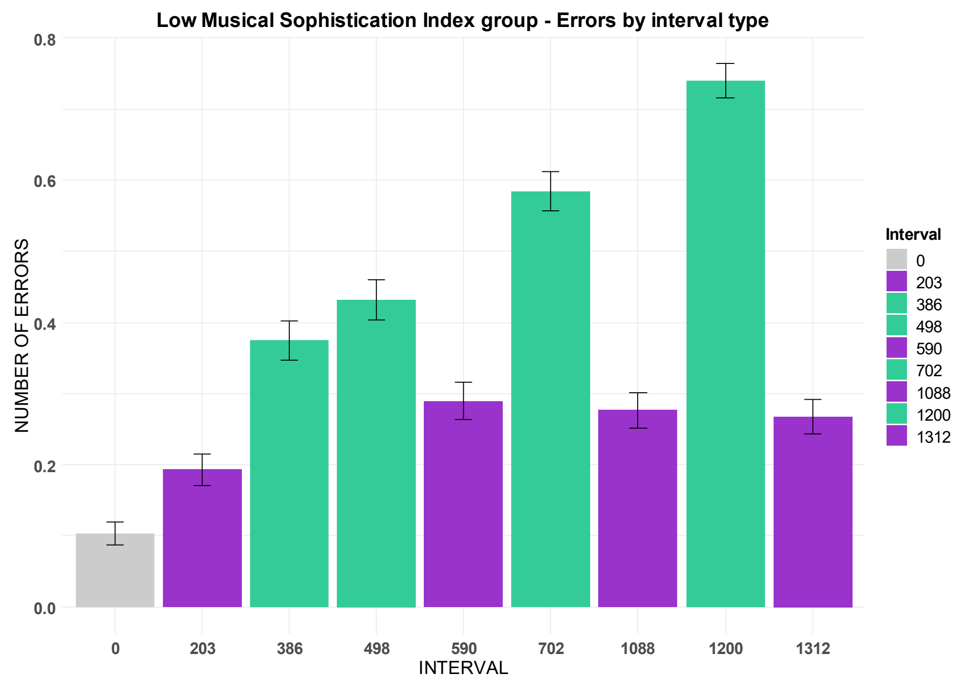


Fig. 6b: Errors by interval type in the first task considering Low Gold-MSI group. Dissonant intervals are represented in violet, consonant intervals in green and unison in grey.

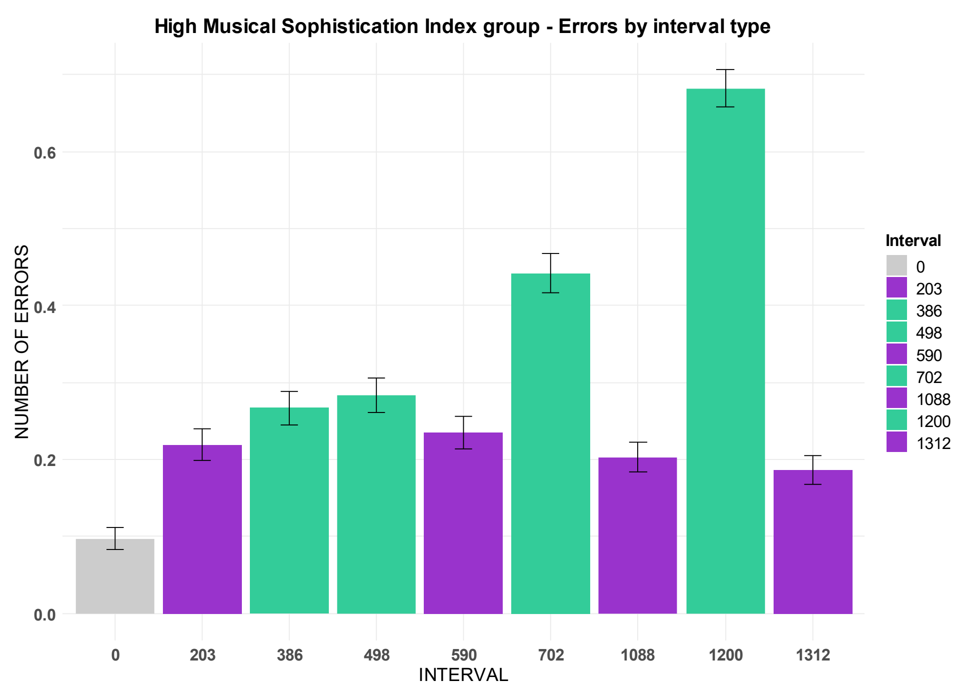


Fig. 6c: Errors by interval type in the first task considering High Gold-MSI group. Dissonant intervals are represented in violet, consonant intervals in green and unison in grey.

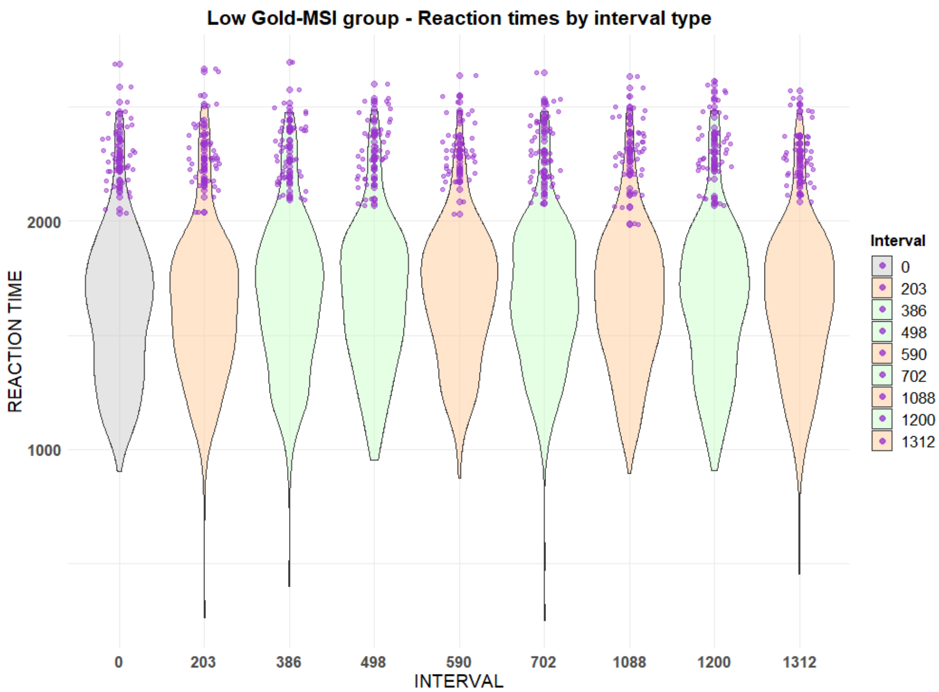


Fig. 6d: Reaction times by interval type in the first task in Low Gold-MSI group. Dissonant intervals are in orange, consonant in green and unison in grey. Dots represent mean values for each participant.

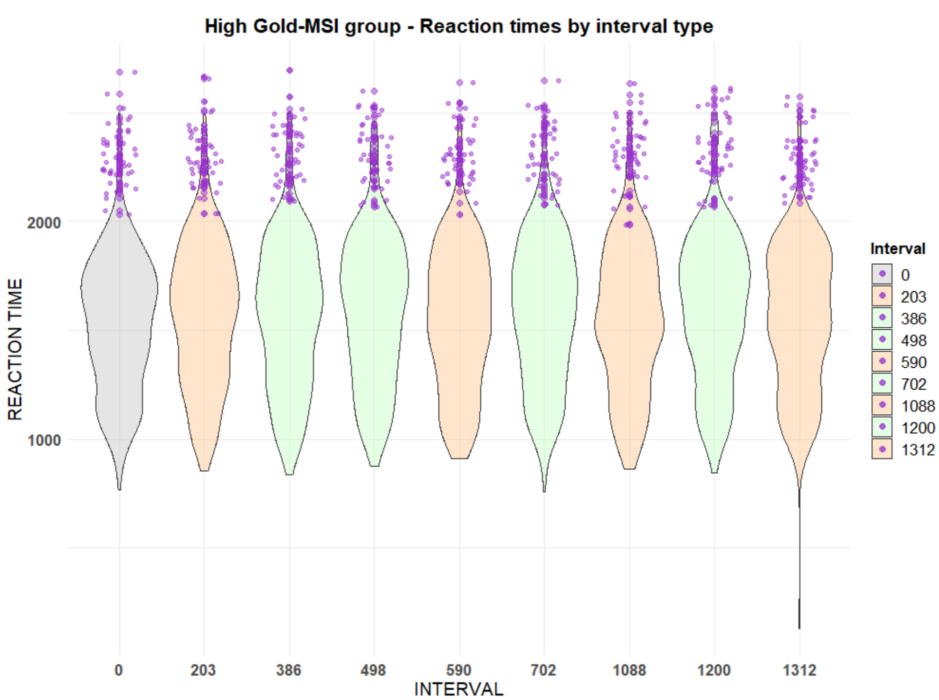


Fig. 6e: Errors by interval type in the first task in High Gold-MSI group. Dissonant intervals are in orange, consonant in green and unison in grey. Dots represent mean values for each participant.

On the fusion judgement task, mean and standard deviations were also calculated separately for each interval, comprehending the full sample, as presented in Table 4a (see also Fig. 7a-b-c). Considering the full sample, mean fusion rating was the highest for octave interval, followed by perfect fifth interval. Considering the two groups High and Low Gold-MSI, ratings for the perfect fourth interval were higher than the ratings for the perfect fifth. The group High Gold-MSI assigned the lowest scorings to major ninth interval.

FUSION RATING	L Group		H Group		TOT	
INTERVAL IN CENTS	MEAN	SD	MEAN	SD	MEAN	SD
0-203	2.13	1.03	2.22	1.14	2.17	1.08
0-386	2.64	1.10	2.45	1.06	2.54	1.08
0-498	2.86	1.12	2.53	1.04	2.69	1.09
0-590	2.54	1.06	2.11	1.10	2.32	1.1
0-702	2.85	1.05	2.27	1.13	2.80	1.09
0-1088	2.39	1.08	2.14	1.08	2.26	1.08
0-1200	3.01	1.15	2.90	1.15	2.96	1.15
0-1312	2.54	1.12	1.93	0.95	2.24	1.08

Table 4a: means and standard deviations calculated by interval for fusion judgement task. In "TOT" column there are the values for the whole sample. Octave interval (highlighted in purple) has the greatest mean rating and standard deviation, followed by perfect fifth interval (highlighted in pink). About the two groups: octave interval (highlighted in purple) has the greatest mean rating and standard deviation, followed in Low Gold-MSI group by perfect fourth interval (highlighted in pink) and perfect fifth interval (lighter pink). In High Gold-MSI group instead of perfect fifth interval, the major third interval (light pink) stands after the perfect fourth interval. Fusion ratings in High Gold-MSI group for major ninth interval (highlighted in yellow) are the lowest ones considering both groups.

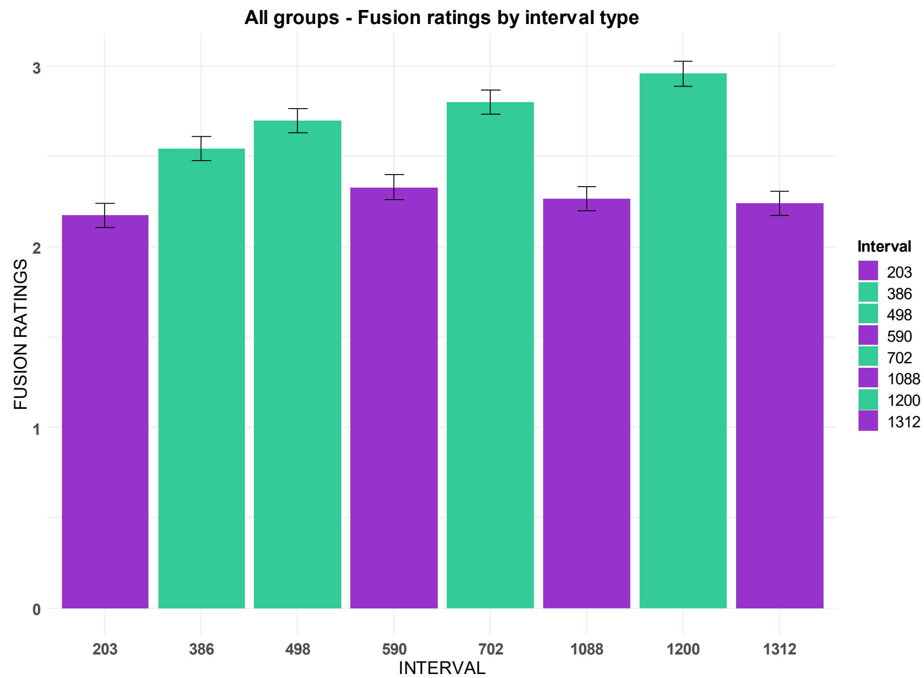
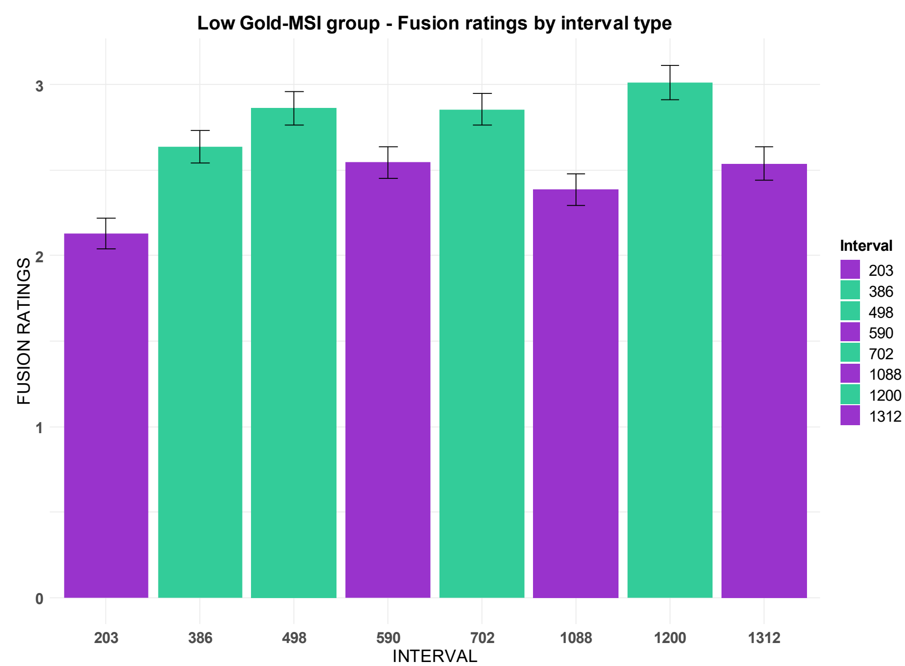


Fig 7a: fusion ratings in the fusion judgement task by interval type, calculated on the sample overall. Dissonant intervals are represented in violet, consonant intervals in green.



.

Fig 7b: fusion ratings in the fusion judgement task by interval type, in Low Gold-MSI group. Dissonant intervals are represented in violet, consonant intervals in green.

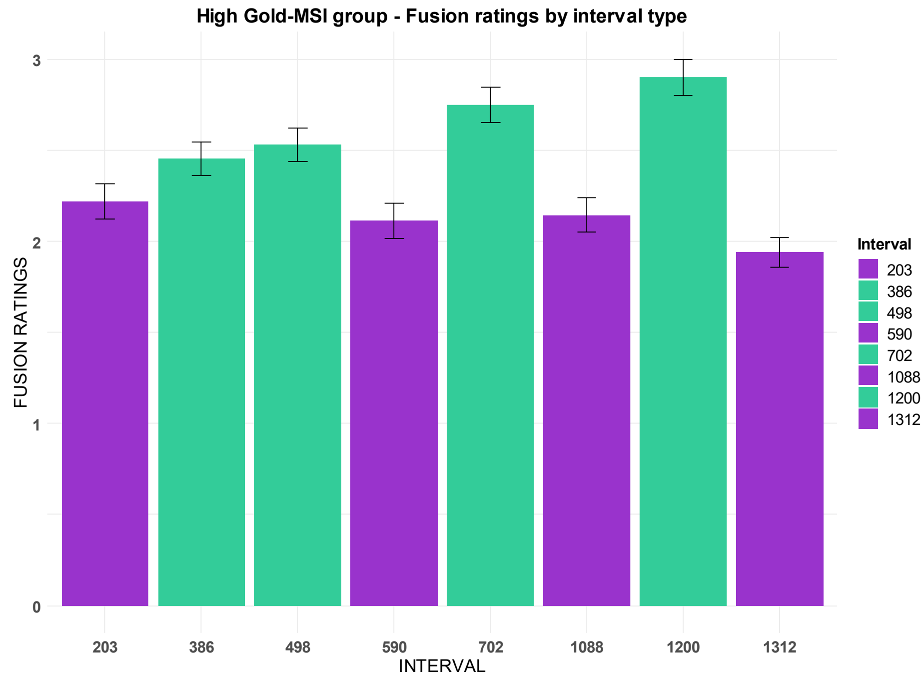


Fig 7c: fusion ratings in the fusion judgement task by interval type, in High Gold-MSI group. Dissonant intervals are represented in violet, consonant intervals in green.

Consonance ratings descriptive statistics are presented in the following Table 4b and Fig. 8a-b-c.

CONSONANCE RATING	L Group		H Group		TOT	
INTERVAL IN CENTS	MEAN	SD	MEAN	SD	MEAN	SD
0-203	2.11	1.01	2.11	1.01	2.11	1.01
0-386	2.46	1.10	2.65	1.01	2.55	1.06
0-498	2.59	1.12	2.80	1.00	2.69	1.06
0-590	2.45	1.04	2.26	1.02	2.35	1.03
0-702	2.55	1.09	3.03	0.97	2.79	1.06
0-1088	2.21	1.03	2.13	0.97	2.17	1.01
0-1200	2.87	1.05	2.96	1.03	2.92	1.04
0-1312	2.20	1.03	1.97	1.09	2.09	1.07

Table 4b: means and standard deviations calculated by interval for consonance judgement task. In "TOT" column there are the values for the whole sample. Octave interval (highlighted in purple) has the greatest mean rating, followed by perfect fifth interval (highlighted in pink). About the two groups: octave

interval (highlighted in purple) has the greatest mean rating in Low Gold-MSI group followed by perfect fourth interval (highlighted in pink) and perfect fifth interval (lighter pink). In H group perfect fifth interval (green) has the greatest rating, whilst the major third interval (light pink) stands after the perfect fourth interval and octave. Consonance ratings in High Gold-MSI group for major ninth interval (highlighted in yellow) are the lowest ones considering both groups.

In the High Gold-MSI group the most extreme values are found. As in fusion ratings (Table 4a) major ninth interval is rated as less preferred, however in consonance ratings we have the highest ratings for perfect fifth interval instead of octave interval.

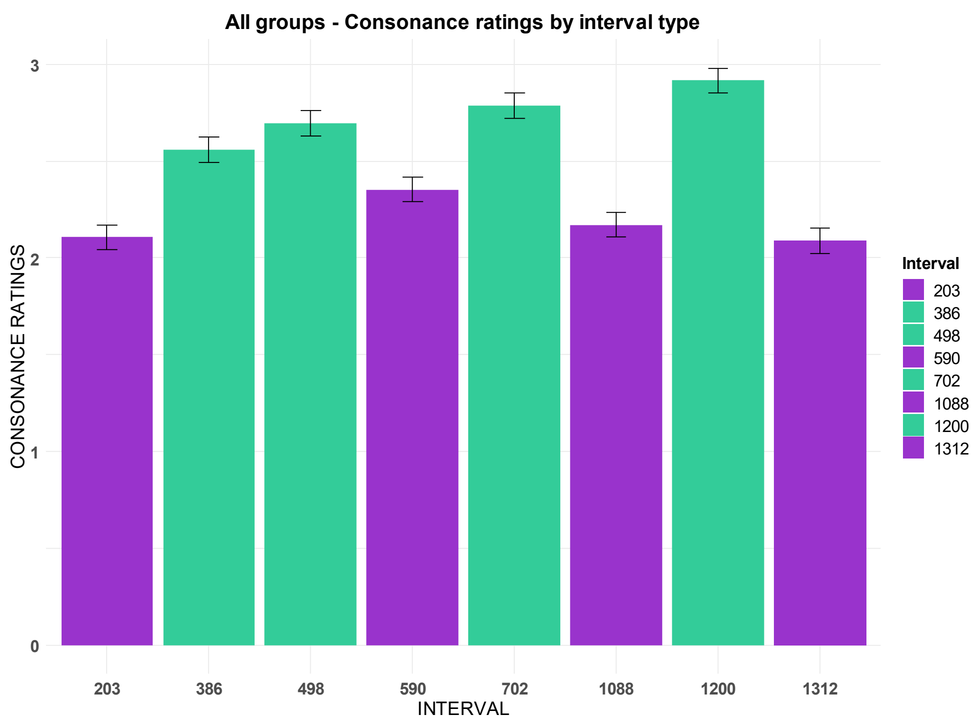


Fig 8a: consonance ratings by interval type, calculated on all the sample. Dissonant intervals are represented in violet, consonant intervals in green.

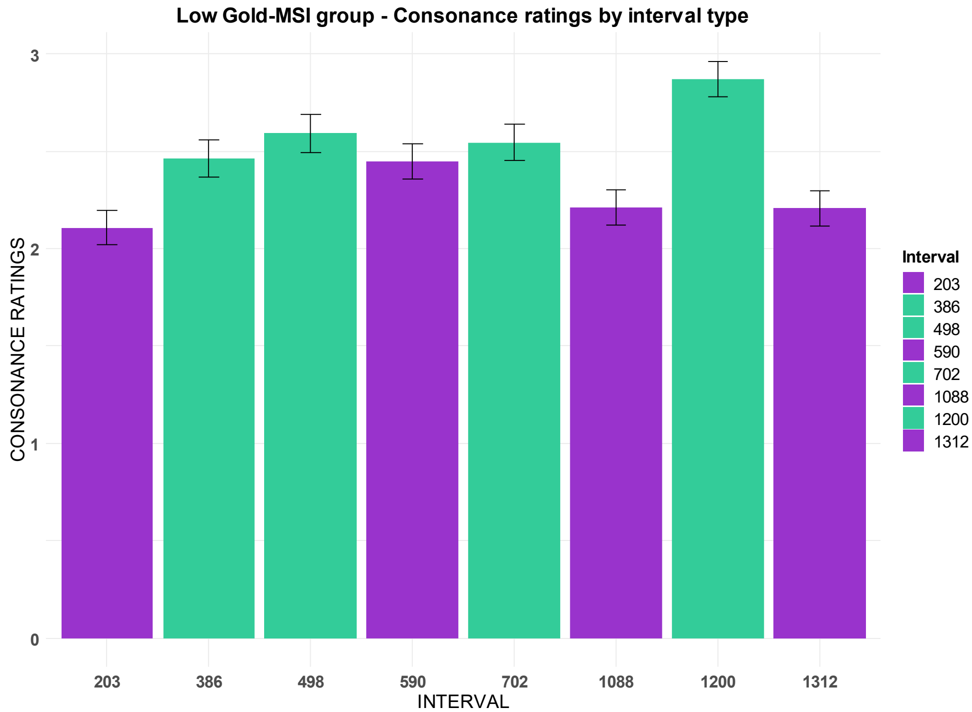


Fig 8b: consonance ratings by interval type, calculated on Low Gold-MSI group. Dissonant intervals are represented in violet, consonant intervals in green.

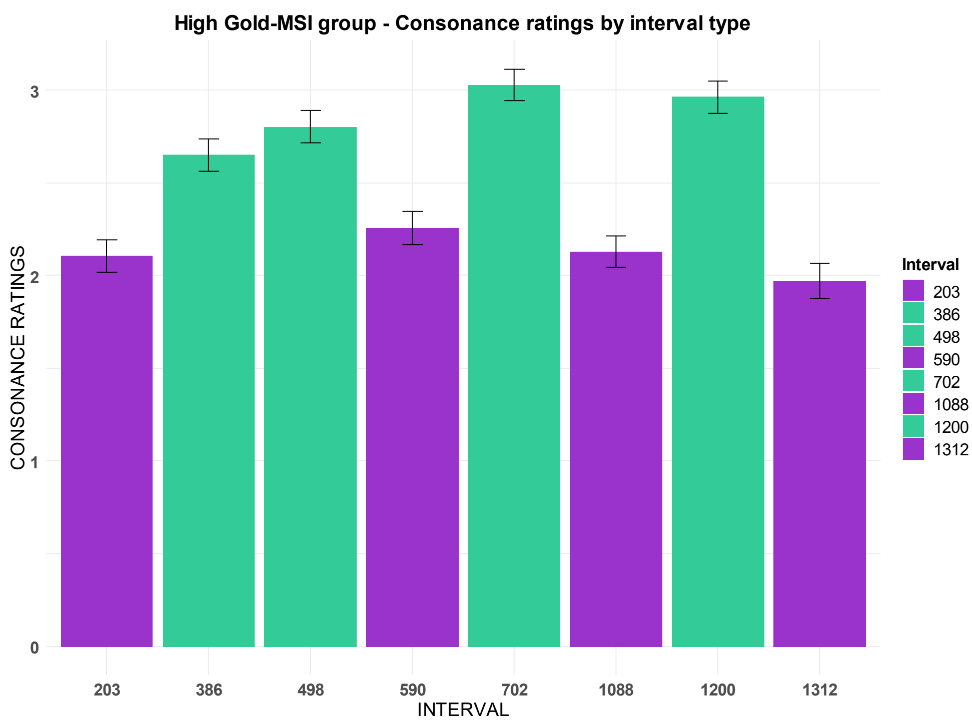


Fig 8c: consonance ratings by interval type, calculated on High Gold-MSI group. Dissonant intervals are represented in violet, consonant intervals in green.

3.2 – The whole sample

On the whole sample, four different analyses with ANOVA one-way repeated measures were performed, with interval type as within-subjects factor. It had 9 levels in the first task, and 8 levels in the fusion judgement and consonance judgement tasks, as the unison interval was not presented in them.

The first ANOVA had as dependent variable the reaction times in the first task. Mauchly's Sphericity test has shown a violation of the assumption, ($W = 0.202$, $p=.001$) so results corrected with Greenhouse-Geisser method ($\varepsilon = 0.719$) are presented here. The effect of interval type on reaction times was not statistically significant ($F_{(8,344)} = 2.123$, $p = 0.054$, $\eta^2=0.011$).

The second ANOVA was performed on number of errors in the first task as dependent variable. Mauchly's Sphericity test has shown a violation of the assumption, ($W = 0.091$, $p =1.63e^{-07}$) so results corrected with Greenhouse-Geisser method ($\varepsilon = 0.607$) are presented here. The effect of interval type on error number was statistically significant ($F(8,344)=56.71$, $p =1.91e^{-36}$, $\eta^2=0.431$).

The third ANOVA was performed on fusion ratings as dependent variable. Mauchly's Sphericity test has shown a violation of the assumption, ($W = 0.054$, $p = 3.11e^{-13}$) so results corrected with Greenhouse-Geisser method ($\varepsilon = 0.473$) are presented here. The effect of interval type on fusion ratings was statistically significant ($F(7,301)=11.58$, $p = 2.89e^{-07}$, $\eta^2=0.156$).

The fourth ANOVA was performed on consonance ratings as dependent variable. Mauchly's Sphericity test has shown a violation of the assumption, ($W = 0.119$, $p=5.02e^{-08}$) so results corrected with Greenhouse-Geisser method ($\varepsilon = 0.622$) are presented here. The effect of interval type on mean consonance ratings was statistically significant ($F(7,301)=24.09$, $p = 4.72e^{-17}$, $\eta^2=0.25$).

Between reaction times and errors in the first two trials, a negligible positive correlation (r = 0.095, p = $7.886e^{-15}$) emerged as significant. The correlation between fusion ratings and consonance ratings showed a similar result (r = -0.046, p = 0.032).

3.3 – High Gold-MSI and Low Gold-MSI groups

On the two groups High Gold-MSI and Low Gold-MSI the four different analyses with ANOVA one-way repeated measures were performed separately, with interval type as within-subjects factor. Interval type had 9 levels in the first task, and 8 levels in the fusion judgement and consonance judgement tasks, as the unison interval was not presented in them. Results are shown in Table 5a-b-c-d.

	ANOVA one-way within (interval type, 9 levels) - Reaction Times						
GROUP	W	p-value	ε	Df	F	p-value	η^2
H	0.03	0.003*	0.568	8-168	2.27	0.059	0.021
L	0.04	0.013*	0.014	8-168	1.01	0.413	0.013

Table 5a: results of ANOVA one-way repeated measures on interval type as factor with nine levels and reaction times as dependent variable, for both groups High Gold-MSI and Low Gold-MSI. Significancy is signaled with an asterisk. From left to right: results of Mauchly's test for Sphericity assumption (W and p-value), Greenhouse-Geisser correction (ε), degrees of freedom for the F statistic (Df), value of the F statistic (F) and corresponding p-value, generalized eta squared (η^2).

	ANOVA one-way within (interval type, 9 levels) - Number of errors						
GROUP	W	p-value	ε	Df	F	p-value	η^2
H	0.01	$3.49e^{-05}$*	0.387	8-168	32.21	$3.80e^{-15}$*	0.387
L	0.04	0.013*	0.571	8-168	27.97	$1.43e^{-16}$*	0.526

Table 5b: results of ANOVA one-way repeated measures on interval type as factor with nine levels and error number as dependent variable, for both groups High Gold-MSI and Low Gold-MSI. Significancy is signaled with an asterisk. From left to right: results of Mauchly's test for Sphericity assumption (W and p-value), Greenhouse-Geisser correction (ε), degrees of freedom for the F statistic (Df), value of the F statistic (F) and corresponding p-value, generalized eta squared (η^2).

	ANOVA one-way within (interval type, 8 levels) – Fusion ratings					
GROUP	W	p-value	ε	Df	F	p-value	η^2
H	0.02	$5.22e^{-05}$*	0.431	7-147	7.64	0.001*	0.930
L	0.04	0.001*	0.442	7-147	5.64	0.001*	0.176

Table 5c: results of ANOVA one-way repeated measures on interval type as factor with eight levels and fusion ratings as dependent variable, for both groups High Gold-MSI and Low Gold-MSI. Significancy is signaled with an asterisk. From left to right: results of Mauchly's test for Sphericity assumption (W and p-value), Greenhouse-Geisser correction (ε), degrees of freedom for the F statistic (Df), value of the F statistic (F) and corresponding p-value, generalized eta squared (η^2).

	ANOVA one-way within (interval type, 8 levels) – Consonance ratings						
GROUP	W	p-value	ε	Df	F	p-value	η^2
H	0.04	0.001*	0.567	7-147	18.21	$9.64e^{-11}$*	0.336
L	0.04	0.001*	0.569	7-147	9.05	$4.14e^{-06}$*	0.199

Table 5d: results of ANOVA one-way repeated measures on interval type as factor with eight levels and consonance ratings as dependent variable, for both groups High Gold-MSI and Low Gold-MSI. Significancy is signaled with an asterisk. From left to right: results of Mauchly's test for Sphericity assumption (W and p-value), Greenhouse-Geisser correction (ε), degrees of freedom for the F statistic (Df), value of the F statistic (F) and corresponding p-value, generalized eta squared (η^2).

Results of ANOVA's analysis on the two groups show a results' pattern like the whole sample. ANOVA on reaction times did not show significance for neither of the two groups, while ANOVAs on error numbers, fusion ratings, and consonance ratings were statistically significant.

Correlation tests between reaction times and number of errors, and between fusion and consonance ratings, were performed separately on the two groups High Gold-MSI and Low Gold-MSI. Results are reported in Table 6a-6b. A significant correlation is observable in group H for both tests, however it is a weak-negligible effect.

Pearson's correlation on reaction times and number of errors					
GROUP	t	p-value	df	int	cor
H	8.882	$2.2e^{-16}$*	3701	0.1127821 0.1758599	0.1444678
L	0.75039	0.4531	2881	-0.02253807 0.05045877	0.01397897

Table 6a: results of correlation tests between reaction times and number of errors in the first task, for both groups High Gold-MSI and Low Gold-MSI. Significancy is signaled with an asterisk. From left to right: value of the test statistic (t and p-value), degrees of freedom (df), 95% confidence interval (int) and correlation value (cor).

Spearman's correlation on fusion ratings and consonance ratings			
GROUP	S	p-value	rho
H	209575695	0.02752*	-0.06782676
L	200309535	0.4454	-0.02351894

Table 6b: results of correlation tests between fusion ratings and consonance ratings, for both groups High Gold-MSI and Low Gold-MSI. Significancy is signaled with an asterisk. From left to right: value of the test statistic (S and p-value), and correlation value (rho).

4 – Results and discussion

From the ANOVA one-way repeated measures between interval type and reaction times, no significant effects emerged, considering the total sample as well as separating the two groups with high and low Musical Sophistication Index. This is consistent with the findings in the first experiment of DeWitt & Crowder (1987), as the results of the second ANOVA, that take into account error number. There was a statistically significant difference between interval types on the number of errors. Octave intervals (1200 cents) had the largest number of errors, as presented in Table 7 and Fig. 6a-b-c, followed by perfect fifth intervals (702 cents) and then perfect fourth intervals (498 cents). The lowest number of errors was observed in major ninth (1312 cents) and major second (203 cents) intervals.

	NUMBER OF ERRORS		
INTERVAL IN CENTS	Low Gold-MSI group	High Gold-MSI group	TOTAL
0-203	62	90	152
0-386	115	108	223
0-498	133	112	245
0-590	88	100	188
0-702	180	171	351
0-1088	89	88	177
0-1200	244	264	508
0-1312	92	79	171
TOT	1038	1054	2092

Table 7: Errors by interval for judgement in the first task (e.g., Do you hear one or two sounds?) for all the groups and the sample overall. Octave interval (highlighted in purple) has the greatest number of errors, followed by perfect fifth interval (highlighted in pink) and perfect fourth interval (highlighted in orange).

Both fusion ratings and consonance ratings analysis shown the effect of interval type. Observing Table 4a and Fig.7a-b-c octave interval shows the highest mean fusion judgement, followed by perfect fifth and perfect fourth intervals. The pattern is visible also in Table 4b and Fig. 8a-b-c, with consonance ratings.

However, comparing Fig.8b and Fig.8c – and values in Table 4b – it is noticeable that consonant ratings for perfect fifth interval are higher than octave interval ratings. Moreover, ratings for major ninth interval are the lowest. Considering mean scoring ranges (Table 8a-b) of both fusion and consonance judgements, we can observe how the High Gold-MSI group seems to show less variability – minimum values in particular – than Low Gold-MSI group. Mean fusion ratings were in general lower for the High Gold-MSI group than the Low Gold-MSI group.

MEAN FUSION RATINGS – HIGH GOLD-MSI GROUP						
INTERVAL IN CENTS	Min	1st Qu.	Median	Mean	3rd Qu.	Max
0-203	1.167	1.667	2	2.220	2.75	3.5
0-386	1.667	2	2.5	2.455	2.833	3.333
0-498	1	2.167	2.5	2.53	3	3.667
0-590	1	1.667	2	2.114	2.625	3
0-702	1	2.208	2.917	2.75	3.292	4
0-1088	1.333	1.667	2.083	2.144	2.667	3.5
0-1200	1	2.667	3.167	2.902	3.5	4
0-1312	1	1.5	1.833	1.939	2.333	3.333
MEAN FUSION RATINGS – LOW GOLD-MSI GROUP						
INTERVAL IN CENTS	Min	1st Qu.	Median	Mean	3rd Qu.	Max
0-203	1	1.708	2.083	2.129	2.5	3.833
0-386	1.833	2.167	2.583	2.636	3	3.833
0-498	2	2.375	2.917	2.864	3.167	3.833
0-590	2	2.208	2.5	2.545	2.792	3.333
0-702	1.5	2.5	2.833	2.856	3.292	3.667
0-1088	1.5	2	2.333	2.386	2.792	3.833
0-1200	1.333	2.292	3.167	3.015	3.625	4
0-1312	1.667	2.042	2.333	2.538	3	4

Table 8a: minimum (Min), first quantile (1st Qu.), median, mean, third quantile (3rd Qu.), and maximum (Max) of mean fusion ratings per participant, for both groups High Gold-MSI and Low Gold-MSI. Dissonant intervals are highlighted in violet.

MEAN CONSONANCE RATINGS – HIGH GOLD-MSI GROUP						
INTERVAL IN CENTS	Min	1st Qu.	Median	Mean	3rd Qu.	Max
0-203	1.167	1.667	1.917	2.106	2.583	3.5
0-386	1.667	2.208	2.5	2.652	3.167	3.667
0-498	2	2.542	2.75	2.803	3.125	3.667
0-590	1.5	1.708	2.25	2.258	2.792	3.333
0-702	1.833	2.833	3	3.03	3.458	3.833
0-1088	1	1.667	2	2.129	2.792	3.167
0-1200	2	2.667	2.833	2.962	3.167	3.833
0-1312	1	1.5	1.833	1.97	2.333	3.5
MEAN CONSONANCE RATINGS – LOW GOLD-MSI GROUP						
INTERVAL IN CENTS	Min	1st Qu.	Median	Mean	3rd Qu.	Max
0-203	1.167	1.667	2.167	2.106	2.458	3.167
0-386	1.5	2.167	2.417	2.462	2.667	3.833
0-498	1.833	2.208	2.583	2.591	2.958	4
0-590	1.667	2.167	2.5	2.447	2.667	3
0-702	1.5	2.208	2.583	2.545	2.833	3.333
0-1088	1.333	2	2.167	2.212	2.500	2.833
0-1200	1.667	2.500	2.917	2.871	3.292	3.667
0-1312	1.667	1.833	2.250	2.205	2.500	3.167

Table 8b: minimum (Min), first quantile (1st Qu.), median, mean, third quantile (3rd Qu.), and maximum (Max) of mean consonance ratings per participant, for both groups High Gold-MSI and Low Gold-MSI. Dissonant intervals are highlighted in violet.

Consonance ratings for consonant intervals for the High Gold-MSI group seem higher than in the Low Gold-MSI group. Vice versa in dissonant intervals, where the Low Gold-MSI group shows higher mean ratings, especially in minimum and first quantile values. Moreover, the values in Table 3b-c show a difference between the two groups in reaction times, whereas observing minimum and median values suggests that the High Gold-MSI group could have been faster than Low Gold-MSI group.

These differences found in the sample of this study between participants with low and high Musical Sophistication Index suggest that music experience could have a relevant role in modulating fusion and consonance judgements.

The general tendence to perceived as fused canonically consonant intervals is consistent with the pattern of results in the previous experiments (DeWitt & Crowder, 1987; McPherson et al., 2020) as well as the overall results in this experiment.

This could support an explanation of tonal fusion as linked to harmonic series, but not its strong influence on consonance perception, which seems to be a more complex phenomenon influenced also by culture (Harrison & Pearce, 2019; McPherson et al., 2020).

4.1 – Limits and Perspectives

Future research could take deeper into account musical experience and investigate fusion perception and consonance with respect to being musicians or non-musicians.

A possible limitation in the current study is numerosity of the sample. The applied tasks' counterbalancing featured four different groups, but all of them completed first tasks where was required to them to being as fast as possible in their responses. This could lead to non-controlled order effects. Enlarging the sample would allow to have more groups, thus controlling other possible order effects.

References

Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(36), 11155–11160. https://doi.org/10.1073/pnas.1505768112

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. http://psychtoolbox.org/[http://color.psych.upenn.edu/brainard/papers/Psychtoolbox.pdf

Cook, N. (2021). Music in a global world. In N. Cook (Ed.), *Music: A Very Short Introduction* (p. 0). Oxford University Press. https://doi.org/10.1093/actrade/9780198726043.003.0006

Dawson, M. (2010). *Musical Consonance*. Dictionary of Cognitive Science - University of Alberta. http://www.bcp.psych.ualberta.ca/~mike/Pearl_Street/Dictionary/contents/M/musicalconson.html

DeWitt, L. A., & Crowder, R. G. (1987). Tonal fusion of consonant musical intervals: The oomph in Stumpf. *Perception & Psychophysics*, *41*(1), 73–84. https://doi.org/10.3758/BF03208216

Di Stefano, N., & Bertolaso, M. (2014). Understanding musical consonance and dissonance: epistemological considerations from a systemic perspective. *Systems*, *2*(4), 566–575. https://doi.org/10.3390/systems2040566

Enciclopedia Treccani. (2023a). *Guido monaco d'Arezzo*. Enciclopedia Treccani. https://www.treccani.it/enciclopedia/guido-monaco-d-arezzo/

Enciclopedia Treccani. (2023b). *Weber, Ernst Heinrich*. Enciclopedia Treccani. https://www.treccani.it/enciclopedia/ernst-heinrich-weber

Gill, K. Z., & Purves, D. (2009). A biological rationale for musical scales. *PLoS ONE*, *4*(12). https://doi.org/10.1371/journal.pone.0008144

Grassi, M., & Soranzo, A. (2009). MLP: A MATLAB toolbox for rapid and reliable auditory threshold estimation. *Behavior Research Methods*, *41*(1), 20–28. https://doi.org/10.3758/BRM.41.1.20

Guernsey, M. (1928). The Role of Consonance and Dissonance in Music. *The American Journal of Psychology*, *40*(2), 173. https://doi.org/10.2307/1414484

Harrison, P. M. C., & Pearce, M. T. (2019). Simultaneous Consonance in Music Perception and Composition. *Psychological Review*, *127*(2), 216–244. https://doi.org/10.1037/rev0000169

Huron, D. (1991). Tonal Consonance versus Tonal Fusion in Polyphonic Sonorities. *Music Perception*, *9*(2), 135–154. https://doi.org/10.2307/40285526

John Marenbon. (2021). *Anicius Manlius Severinus Boethius*. Stanford Encyclopedia of Philosophy (Winter 2021 Edition). https://plato.stanford.edu/entries/boethius/

Kleiner, M., Brainard, D., & Pelli, D. (2007). *"What's new in Psychtoolbox-3?" Perception 36 ECVP Abstract Supplement.* http://psychtoolbox.org/[http://www.perceptionweb.com/abstract.cgi?id=v070821

Martinelli, R. (2014). Melting Musics, Fusing Sounds - Stumpf, Hornbostel, and Comparative Musicology in Berlin. In R. Bod, J. Maat, & T. Weststeijn (Eds.), *The Making of the Humanities, Vol.III* (pp. 391–401). Amsterdam University Press. http://journal.oraltradition.org/files/articles/2i/15_saffar.pdf

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*(7613), 547–550. https://doi.org/10.1038/nature18635

McLachlan, N., Marco, D., Light, M., & Wilson, S. (2013). Consonance and pitch. *Journal of Experimental Psychology: General*, *142*(4), 1142–1158. https://doi.org/10.1037/a0030830

McPherson, M. J., Dolan, S. E., Durango, A., Ossandon, T., Valdés, J., Undurraga, E. A., Jacoby, N., Godoy, R. A., & McDermott, J. H. (2020). Perceptual fusion of musical notes by native Amazonians suggests universal representations of musical intervals. *Nature Communications*, *11*(1), 1–14. https://doi.org/10.1038/s41467-020-16448-6

Metfessel, M. (1926). Sonance as a form of tonal fusion. *Psychological Review*, *33*(6), 459–466. https://doi.org/10.1037/h0070587

Müllensiefen, D., Gingras, B., Stewart, L., & Musil, J. J. (2013). Goldsmiths Musical Sophistication Index (Gold-MSI) v1.0 Technical Report and Documentation Revision 0.3. *Technical Report*, 1–69.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. http://www.psych.nyu.edu/pelli/pubs/pelli1997videotoolbox.pdf

Posit team. (2023). *RStudio: Integrated Development Environment for R* (2023.9.1.494). Posit Software, PBC. http://www.posit.co/

Qualtrics. (2020). *Qualtrics* (No. 102023). Qualtrics. https://www.qualtrics.com/

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.r-project.org/

Sam M.S., N. (2013). *Tonal Fusion*. PsychologyDictionary.Org. https://psychologydictionary.org/tonal-fusion/

Schneider, A. (1997). Verschmelzung, "Tonal fusion, and consonance: Carl Stumpf revisited.". In M. Leman (Ed.), *Music, gestalt, and comput- ing: Studies in cognitive and systematic musicology* (pp. 115–143). Springer. http://dx.doi.org/10.1007/ BFb0034111

Sears, D. (2008). Processing Surprise Tension in Tonal Melodies. *Inquiry: The University of Arkansas Undergraduate Research Journal*, *9*(1).

Terhardt, E. (1984). The Concept of Musical Consonance : A Link between Music and Psychoacoustics Author (s): Ernst Terhardt Source : Music Perception : An Interdisciplinary Journal , Vol . 1 , No . 3 , Dedicated to Published by : University of California Press Stable URL : *Music Perception*, *1*(3), 276–295.

The MathWorks Inc. (2023). *MATLAB* (23.2.0.2365128 (R2023b)). The MathWorks Inc.

Wendt, G., & Bader, R. (2019). Analysis and Perception of Javanese Gamelan Tunings. In R. Bader (Ed.), *Computational Phonogram Acrhiving (Current Research in Systematic Musicology)* (pp. 129–144). Springer Nature Switzerland AG 2019.

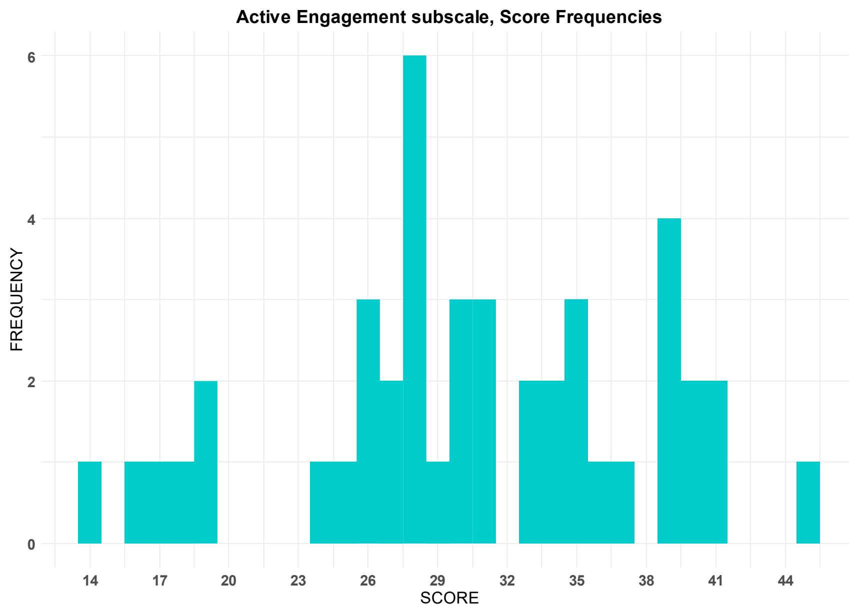
Wood, A., & Lindsay, R. B. (1957). The Physics of Music. *Physics Today*, *10*(4), 36. https://doi.org/10.1063/1.3060334

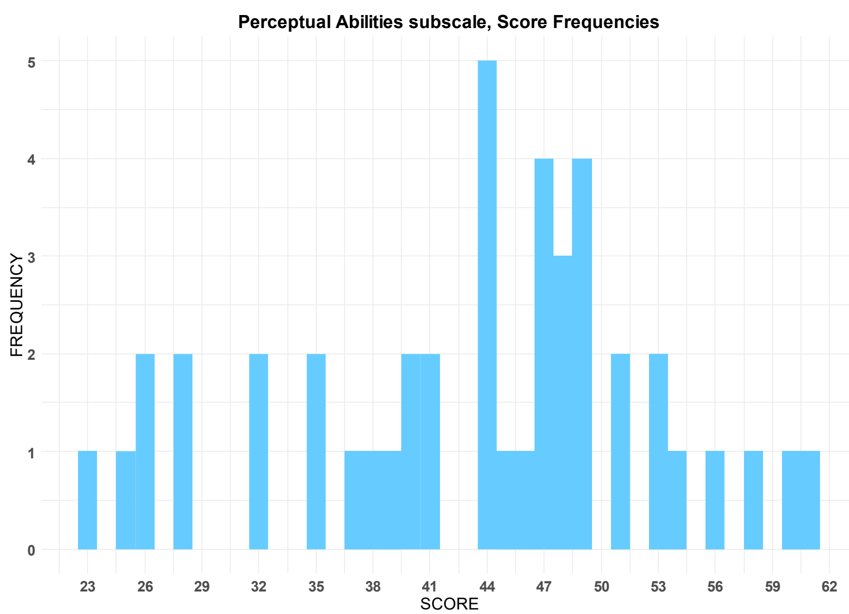
Yost, W. (2015). Psychoacoustics: A Brief Historical Overview. *Acoust. Today*, *11*(3), 46–53.

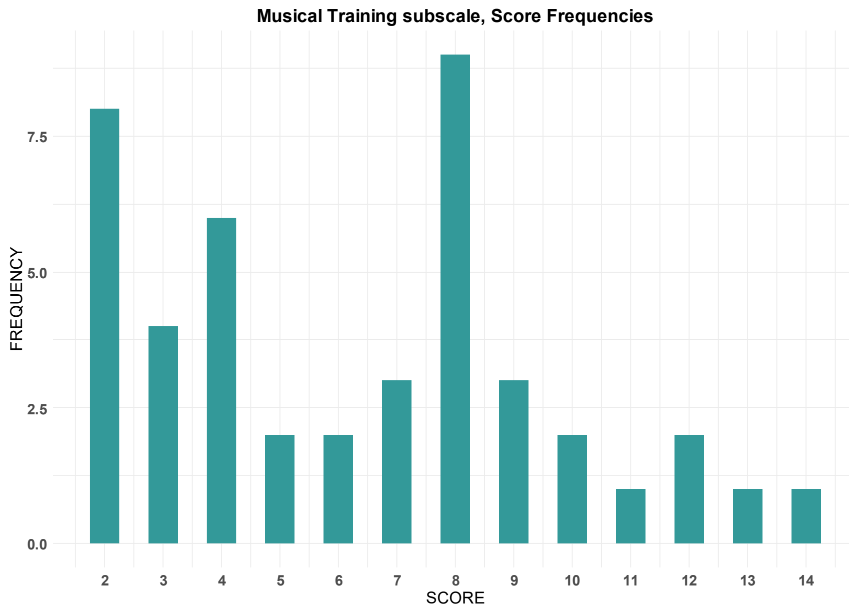
Yuan, J. (2020). *Comparison of Dissonance Distributions Between Timbres of Different Instruments*. *402*(Ssphe 2019), 9–13. https://doi.org/10.2991/assehr.k.200205.003

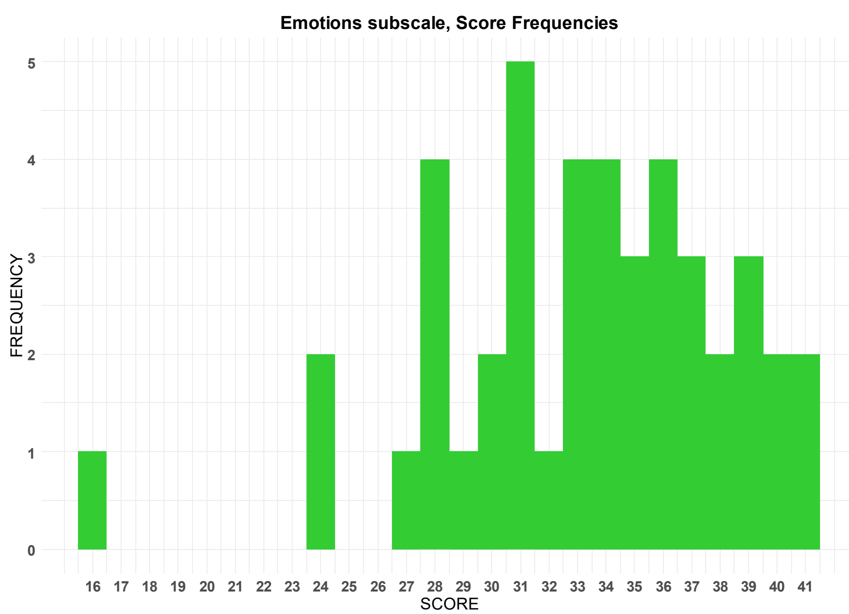
Appendices

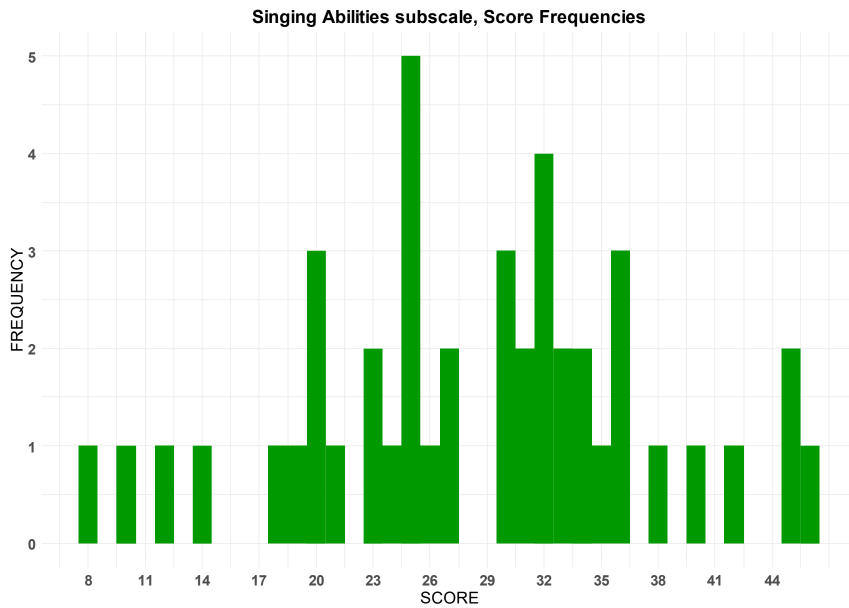
A – Gold-MSI subscales scoring distributions






Musical Training subscale, Score Frequencies
FREQUENCY
0.0
2.5
5.0
7.5
2
3
4
5
6
7
8
9
10
11
12
13
14
SCORE


Emotions subscale, Score Frequencies
FREQUENCY
0
1
2
3
4
5
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
SCORE


Singing Abilities subscale, Score Frequencies
FREQUENCY
0
1
2
3
4
5
8
11
14
17
20
23
26
29
32
35
38
41
44
SCORE

B – Italian translation of the Gold-MSI

ORDER	ITEM	ENGLISH	ITALIAN
1	AE_01	*I spend a lot of my free time doing music-related activities.*	Trascorro molto del mio tempo libero facendo attività correlate alla musica
2	EM_01	*I sometimes choose music that can trigger shivers down my spine.*	A volte scelgo musica che mi fa venire i brividi lungo la schiena
3	AE_02	*I enjoy writing about music, for example on blogs and forums.*	Mi piace scrivere di musica, per esempio su blog e forum
4	SA_01	*If somebody starts singing a song I don't know, I can usually join in.*	Se qualcuno inizia a cantare una canzone che non conosco, generalmente riesco a partecipare
5	PA_01	*I am able to judge whether someone is a good singer or not.*	Sono capace di giudicare se qualcuno è un bravo cantante o no
6	PA_02	*I usually know when I'm hearing a song for the first time.*	Generalmente so quando sto sentendo una canzone per la prima volta
7	SA_02	*I can sing or play music from memory.*	Sono capace di suonare o cantare a memoria
8	AE_03	*I'm intrigued by musical styles I'm not familiar with and want to find out more.*	Sono incuriosito/a da generi musicali che non mi sono familiari e cerco di saperne di più
9	EM_02	*Pieces of music rarely evoke emotions for me.*	I brani musicali raramente provocano emozioni in me
10	SA_03	*I am able to hit the right notes when I sing along with a recording.*	Sono capace di azzeccare le note giuste quando canto insieme ad una registrazione
11	PA_03	*I find it difficult to spot mistakes in a performance of a song even if I know the tune.*	Trovo difficile individuare gli errori nell'interpretazione di una canzone anche se conosco la melodia
12	PA_04	*I can compare and discuss differences between two performances or versions of the same piece of music.*	Posso confrontare e discutere su due esecuzioni o versioni dello stesso brano musicale
13	PA_05	*I have trouble recognizing a familiar song when played in a different way or by a different performer.*	Ho problemi a riconoscere una canzone familiare se suonata in diversi modi o da un diverso esecutore
14	MT_03	*I have never been complimented for my talents as a musical performer.*	Non mi hanno mai fatto i complimenti per il mio talento come musicista

15	AE_05	*I often read or search the internet for things related to music.*	Ho spesso letto o ricercato su internet cose correlate alla musica
16	EM_03	*I often pick certain music to motivate or excite me.*	Spesso scelgo un certo tipo di musica per motivarmi o attivarmi
17	SA_04	*I am not able to sing in harmony when somebody is singing a familiar tune.*	Non sono capace di cantare in armonia quando qualcuno canta una melodia familiare
18	PA_06	*I can tell when people sing or play out of time with the beat.*	Riesco a dire quando le persone cantano o suonano fuori tempo
19	EM_04	*I am able to identify what is special about a given musical piece.*	Sono in grado di identificare cosa c'è di speciale in un determinato brano musicale
20	EM_05	*I am able to talk about the emotions that a piece of music evokes for me.*	Sono capace di parlare delle emozioni che un brano musicale provoca in me
21	AE_06	*I don't spend much of my disposable income on music.*	Non spendo molto del mio reddito per la musica
22	PA_07	*I can tell when people sing or play out of tune.*	So dire se le persone sono stonate quando cantano o suonano
23	PA_08	*When I sing, I have no idea whether I'm in tune or not.*	Quando canto, non ho idea se sono intonato/a oppure no
24	AE_07	*Music is kind of an addiction for me - I couldn't live without it.*	La musica è una specie di dipendenza per me; non potrei vivere senza
25	SA_05	*I don't like singing in public because I'm afraid that I would sing wrong notes.*	Non mi piace cantare in pubblico perché ho paura di cantare note sbagliate
26	PA_09	*When I hear a piece of music I can usually identify its genre.*	Quando ascolto un brano generalmente riesco a identificarne il genere
27	MT_07	*I would not consider myself a musician.*	Non mi considererei un musicista
28	AE_09	*I keep track of new music that I come across (e.g. new artists or recordings).*	Prendo nota della nuova musica che incontro (es. nuovi artisti o registrazioni)
29	SA_06	*After hearing a new song two or three times, I can usually sing it by myself.*	Dopo aver ascoltato una canzone per due o tre volte, in genere riesco a cantarla autonomamente
30	SA_07	*I only need to hear a new tune once and I can sing it back hours later.*	Ho bisogno di ascoltare una melodia solo una volta e posso tornare a cantarla ore dopo
31	EM_06	*Music can evoke my memories of past people and places.*	La musica può rievocarmi dei ricordi di persone e luoghi del passato

32	MT_01	*I engaged in regular, daily practice of a musical instrument (including voice) for_ years.*	Sono stato/a impegnato/a quotidianamente e regolarmente nella pratica di uno strumento musicale (o della voce) per il seguente numero di anni:
33	MT_02	*At the peak of my interest, I practised my primary instrument for _ hours per day.*	All'apice del mio interesse, mi sono esercitato/a con il mio strumento primario (o con la voce) per il seguente numero di ore al giorno:
34	AE_04	*I have attended _ live music events as an audience member in the past twelve months.*	Ho frequentato il seguente numero di eventi di musica dal vivo come membro del pubblico negli ultimi dodici mesi:
35	MT_04	*I have had formal training in music theory for _ years.*	Ho ricevuto lezioni formali di teoria della musica per il seguente numero di anni:
36	MT_05	*I have had _ years of formal training on a musical instrument (including voice) during my lifetime.*	Ho ricevuto lezioni formali di uno strumento musicale (o di canto) durante la mia vita per il seguente numero di anni:
37	MT_06	*I can play _ musical instruments.*	So suonare il seguente numero di strumenti musicali:
38	AE_08	*I listen attentively to music for _ per day.*	Ascolto attentamente musica per la seguente quantità di tempo al giorno:
39	BI_01	*The instrument I play best (including voice) is:*	Lo strumento che suono meglio (inclusa la voce) è:
40	ST_01	*What age did you start to play an instrument?*	A quale età hai iniziato a suonare uno strumento?
41	AP_01	*Do you have absolute pitch? Absolute or perfect pitch is the ability to recognise and name an isolated musical tone without a reference tone, e.g. being able to say 'F#' if someone plays that note on the piano.*	Hai l'orecchio assoluto? L'orecchio assoluto è la capacità di riconoscere e nominare una nota musicale isolata senza note di riferimento, ad es. essere in grado di dire 'Fa#' se qualcuno suona quella nota al pianoforte.

answer	**English**	**Italian**
1	Completely disagree	Completamente in disaccordo
2	Strongly disagree	Molto in disaccordo
3	Disagree	In disaccordo
4	Neither agree or disagree	Né in disaccordo né d'accordo
5	Agree	D'accordo

6	Strongly agree	Molto d'accordo
7	Completely agree	Completamente d'accordo
AE_04	**AE_04**	**AE_04**
	0	0
	1	1
	2	2
	3	3
	4-6	4-6
	7-10	7-10
	11 or more	11 o più
AE_08	**AE_08**	**AE_08**
1	0-15 min	0-15 min
2	15-30 min	15-30 min
3	30-60 min	30-60 min
4	60-90 min	60-90 min
5	2 hrs	2 ore
6	2-3 hrs	2-3 ore
7	4 hrs or more	4 ore o più
MT_01	**MT_01**	**MT_01**
1	0	0
2	1	1
3	2	2
4	3	3
5	4-5	4-5
6	6-9	6-9
7	10 or more	10 o più
MT_02	**MT_02**	**MT_02**
1	0	0
2	0.5	0.5
3	1	1
4	1.5	1.5
5	2	2
6	3-4	3-4
7	5 or more	5 o più
MT_04	**MT_04**	**MT_04**
1	0	0
2	0.5	0.5
3	1	1
4	2	2
5	3	3
6	4-6	4-6
7	7 or more	7 o più
MT_05	**MT_05**	**MT_05**

1	0	0
2	0,5	0,5
3	1	1
4	2	2
5	3-5	3-5
6	6-9	6-9
7	10 or more	10 o più
MT_06	**MT_06**	**MT_06**
1	0	0
2	1	1
3	2	2
4	3	3
5	4	4
6	5	5
7	6 or more	6 o più
ST_01	**ST_01**	**ST_01**
	2-19	2-19
	I don't play any instrument.	Non suono nessuno strumento
	2 years or younger	2 anni o meno
	19 years or older	19 anni o più
AP_01	**AP_01**	**AP_01**
1	Yes	Si
2	No	No
BI_01	**BI_01**	**BI_01**
NO	I don't play any instrument.	Non suono nessuno strumento
voic	Voice	Voce
pian	Piano	Pianoforte
guit	Guitar	Chitarra
drum	Drums	Percussioni
xylo	Xylophone	Xilofono
flut	Flute	Flauto
oboe	Oboe	Oboe
clar	Clarinet	Clarinetto
baso	Bassoon	Fagotto
trum	Trumpet	Tromba
trom	Trombone	Trombone
tuba	Tuba	Tuba
saxo	Saxophone	Sassofono
bugl	Bugle	Corno
viol	Violin	Violino
cell	Cello	Violoncello
vioa	Viola	Viola
bass	Double bass	Contrabbasso
harp	Harp	Arpa

_other	Other	Altro

Instructions	**English**	**Italian**
instr	Please select the most appropriate category:	Per favore selezioni la categoria più appropriata:

C – Hardware and software

The computer had an Intel(R) Core (TM) i5-7400 CPU 3.00GHz processor, with 16GB RAM, and 64-bit operative system. Product ID: 00330-50157-12329-AAOEM. Operative system: edition Windows 10 Pro, 22H2 version, installed 20/01/2023, build 19045.3570, Windows Feature Experience Pack 1000.19052.1000.0.

The audio card was a Focusrite Scarlett 4i4 (3rd Gen), Sample Rate: 48000, Buffer Size:192, Clock: Internal (SYNCED). As over-head earphones were used Sennheiser HDA 300, frequencies interval 20 Hz - 20 kHz, sound pressure SPL (dB) 117.

The display was a DELL U2715H Monitor LCD 2560x1440 Pixel Resolution 27" LCD Display HDMI, MHL, DP, Mini DP Connectivity, 4 Port USB Hub & 1 BC Port. Responses were collected using a Compaq Prolinea keyboard, FCC ID EW4KPQ2479, model KPQ-E99YC, input 5.0 VDC.

The main experiment was written in MATLAB (The MathWorks Inc., 2023) using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). The audio stimuli were created using the MLP toolbox (Grassi & Soranzo, 2009).

Statistical analyses were conducted using R (R Core Team, 2021) and RStudio (Posit team, 2023).

D – MATLAB scripts

Function used in the first two tasks

```
function fusion_rt(nsub, subname, subsex, subage, order)

%% psychtoobox settings
Screen('Preference', 'SkipSyncTests', 1);
Screen('Preference','SuppressAllWarnings');
screenNumber=max(Screen('Screens'));
black=BlackIndex(screenNumber);
white=WhiteIndex(screenNumber);

rect1=Screen('Rect', screenNumber);
[w, ~]=Screen('OpenWindow',screenNumber, black, rect1);
%[w, ~]=Screen('OpenWindow',screenNumber, black, [0, 0, 400, 400]);
ifi=Screen('GetFlipInterval', w);
slack=ifi/2;
prioritylevel=MaxPriority(w);
rand('state', sum(100*clock));
Screen('TextFont', w, 'Lucida Sans');
Screen('TextSize', w, 24);
Screen('TextStyle', w, 1);
KbName('UnifyKeyNames');
mKey=KbName('m'); %KbName 77
zKey=KbName('z'); %KbName 90
% sound settings
sf = 48000;
d = 2000; % duration of the stimulus
InitializePsychSound;
pahandle = PsychPortAudio('Open', [], [], 0, sf, 1);

%% EXPERIMENT SETTINGS
ISI = [0.5, 1];
% each stimulus is repeated 10 times
repetitions = 1;
    intervals = [0, 203, 386, 498, 590, 702, 1088, 1200, 1312]; % size of
the intervals in cents
    f0 = [110, 110*2^(2/12), 220, 220*2^(2/12), 440, 440*2^(2/12)];

% event table that includes the characteristics of all the experiment
trials
[event_table, ntrials]=makeeventtable(intervals, f0, ISI, repetitions);
trialorder=randperm(ntrials);
% trialorder=1:ntrials;

% the "output" variable is used later to store all the data of the
% experiment trial by trial
output=zeros(ntrials, 7);

%% experiment
```

```
WaitSecs(3);
HideCursor
Screen('Flip', w);
% main cicle of the experiment with all the trials
for trial=1:ntrials
    WaitSecs(1);
    if trial== ntrials/2
        DrawFormattedText(w, 'Premere la barra spaziatrice per continuare',
'center', 'center', white);
        Screen('Flip', w);
        while 1
            [~, ~, keyCode] = KbCheck;
            if keyCode(KbName('space'))
                break;
            end
        end
        WaitSecs(1);
    end
    % generate the string for the current trial
    s = SynthesizeSound(sf, d, event_table(trialorder(trial), 1),
event_table(trialorder(trial), 2));
    % show the fixation point
    DrawFormattedText(w, '+', 'center', 'center', white);
    PsychPortAudio('FillBuffer', pahandle, s');
    timeElapsed = 0;
    % show the trial at maximal priority
    Priority(prioritylevel);
    % keep the fixation for 180 ms
    t=Screen('Flip', w);
    t0=Screen('Flip', w, t+0.5-slack);
    PsychPortAudio('Start', pahandle, [], t0+event_table(trialorder(trial),
3));
    % collect response
    [~, rispRT, kbCode] = KbCheck;
    while kbCode(zKey)==0 && kbCode(mKey)==0 && timeElapsed < 3.0
        [~, rispRT, kbCode] = KbCheck;
        timeElapsed = rispRT - t0-event_table(trialorder(trial), 3);
    end
    Priority(0);
    PsychPortAudio('Stop', pahandle);
    % store results
    output(trial, 1)=trial;
    output(trial, 2)=event_table(trialorder(trial), 1);% interval
    output(trial, 3)=event_table(trialorder(trial), 2);% f0
    if find(kbCode)
        responses = find(kbCode);
        output(trial, 4)=responses(1);
    else
        output(trial, 4)=999;
    end
    WaitSecs(3);
    if order == "AB"
        % if order = 'AB' then Z unison and M interval responses
```

```
        if kbCode(zKey) && event_table(trialorder(trial), 1) == 0
            rispAC = 1;
        elseif kbCode(mKey) && event_table(trialorder(trial), 1) ~= 0
            rispAC = 1;
        else
            rispAC = 0;
        end
    elseif order == "BA"
        % if order = 'BA' then M unison and Z interval responses
        if kbCode(mKey) && event_table(trialorder(trial), 1) == 0
            rispAC = 1;
        elseif kbCode(zKey) && event_table(trialorder(trial), 1) ~= 0
            rispAC = 1;
        else
            rispAC = 0;
        end
    end
    output(trial, 5)= rispAC;
    output(trial, 6)= 1000*timeElapsed;
    output(trial, 7)= event_table(trialorder(trial), 3);
end
%% end of experiment
if nsub==0
    DrawFormattedText(w, 'La prova è finita', 'center', 'center', white);
else
    DrawFormattedText(w, 'L''esperimento è finito', 'center', 'center',
white);
end
Screen('Flip', w);
WaitSecs(3);
Screen('Flip', w);
ShowCursor;
Screen('CloseAll');
PsychPortAudio('Close', pahandle);

%% write datafile
% check whether the datafile exists in order to write if necessary the
% names of the variables
time=clock;
time=time(4:5);
time=[num2str(time(1)),':',num2str(time(2))];
if nsub ~= 0
    if ~exist('fusion_rt.txt', 'file')
        datafile = fopen('fusion_rt.txt', 'a');
        fprintf (datafile,
'subname\tsubsex\tsubage\tnsub\tdate\ttime\ttrial\tinterval\tf0\trispSE\tri
spAC\trispRT\tISI\tOrder\n');
    else
        datafile = fopen('fusion_rt.txt', 'a');
    end
    % write all output values
    [rows, columns]=size(output);
    for i=1:rows
```

```
        fprintf(datafile, '%s\t%s\t%2.0f\t%2.0f\t%s\t%s\t', subname,
subsex, subage, nsub, date, time);
        for j=1:columns
            fprintf(datafile, '%4.1f\t', output(i, j));
        end
        fprintf(datafile, '%s\t', order);
        fprintf(datafile, '\n');
    end
    % close the file
    status = fclose (datafile);
end
```

Function used in the fusion rating task

```
function fusion_rating(nsub, subname, subsex, subage, order)

%% psychtoobox settings
Screen('Preference', 'SkipSyncTests', 1);
Screen('Preference','SuppressAllWarnings');
screenNumber=max(Screen('Screens'));
black=BlackIndex(screenNumber);
white=WhiteIndex(screenNumber);
gray=round((white+black)/2);
% This makes sure that on floating point framebuffers we still get a
% well defined gray. It isn't strictly neccessary in this demo:
if gray==white
    gray=white/2;
end
rect1=Screen('Rect', screenNumber);
[w, rect]=Screen('OpenWindow',screenNumber, black, rect1);
%[w, ~]=Screen('OpenWindow',screenNumber, black, [0, 0, 400, 400]);
ifi=Screen('GetFlipInterval', w);
slack=ifi/2;
prioritylevel=MaxPriority(w);
rand('state', sum(100*clock));
Screen('TextFont', w, 'Lucida Sans');
Screen('TextSize', w, 24);
Screen('TextStyle', w, 1);
KbName('UnifyKeyNames');
oneKey=KbName('1!');
twoKey=KbName('2@');
% sound settings
sf = 48000;
d = 2000; % duration of the stimulus
InitializePsychSound;
pahandle = PsychPortAudio('Open', [], [], 0, sf, 1);

%% EXPERIMENT SETTINGS
ISI = [0.5, 1];
% each stimulus is repeated 10 times
```

```
repetitions = 1;
    intervals = [203, 386, 498, 590, 702, 1088, 1200, 1312]; % size of the
intervals in cents
    f0 = [110, 110*2^(2/12), 220, 220*2^(2/12), 440, 440*2^(2/12)];

% here I create an event table that includes the characteristics of all the
% trials of the experiment
[event_table, ntrials]=makeeventtable(intervals, f0, ISI, repetitions);
trialorder=randperm(ntrials);
% trialorder=1:ntrials;

% the "output" variable is used later to store all the data of the
% experiment trial by trial
output=zeros(ntrials, 7);

%% experiment
WaitSecs(3);
HideCursor
Screen('Flip', w);
% main cicle of the experiment with all the trials
for trial=1:ntrials
    WaitSecs(1);
    if trial== ntrials/2
        DrawFormattedText(w, 'Premere la barra spaziatrice per continuare',
'center', 'center', white);
        Screen('Flip', w);
        while 1
            [~, ~, keyCode] = KbCheck;
            if keyCode(KbName('space'))
                break;
            end
        end
        WaitSecs(1);
    end
    % generate the string for the current trial
    s = SynthesizeSound(sf, d, event_table(trialorder(trial), 1),
event_table(trialorder(trial), 2));
    % show the fixation point
    DrawFormattedText(w, '+', 'center', 'center', white);
    PsychPortAudio('FillBuffer', pahandle, s');
    timeElapsed = 0;
    % show the trial at maximal priority
    Priority(prioritylevel);
    % keep the fixation for 180 ms
    t=Screen('Flip', w);
    t0=Screen('Flip', w, t+0.5-slack);
    PsychPortAudio('Start', pahandle, [], t0+event_table(trialorder(trial),
3));
    % collect response
    [~, rispRT, kbCode] = KbCheck;
    while kbCode(oneKey)==0 && kbCode(twoKey)==0
        [~, rispRT, kbCode] = KbCheck;
        timeElapsed = rispRT - t0-event_table(trialorder(trial), 3);
```

```
    end
    Priority(0);
    PsychPortAudio('Stop', pahandle);
    % store results
    output(trial, 1)=trial;
    output(trial, 2)=event_table(trialorder(trial), 1);% interval
    output(trial, 3)=event_table(trialorder(trial), 2);% f0
    if find(kbCode)
        responses = find(kbCode);
        output(trial, 4)=responses(1);
    else
        output(trial, 4)=999;
    end
    WaitSecs(3);

    if kbCode(oneKey) && event_table(trialorder(trial), 1) == 0
        rispAC = 1;
    elseif kbCode(twoKey) && event_table(trialorder(trial), 1) ~= 0
        rispAC = 1;
    else
        rispAC = 0;
    end

    output(trial, 5)= rispAC;
    output(trial, 6)= 1000*timeElapsed;
    output(trial, 7)= event_table(trialorder(trial), 3);
end
%% end of experiment
if nsub==0
    DrawFormattedText(w, 'La prova è finita', 'center', 'center', white);
else
    DrawFormattedText(w, 'L''esperimento è finito', 'center', 'center',
white);
end
Screen('Flip', w);
WaitSecs(3);
Screen('Flip', w);
ShowCursor;
Screen('CloseAll');
PsychPortAudio('Close', pahandle);

%% write datafile
% check whether the datafile exists in order to write if necessary the
% names of the variables
time=clock;
time=time(4:5);
time=[num2str(time(1)),':',num2str(time(2))];
if nsub ~= 0
    if ~exist('fusion_rt.txt', 'file')
        datafile = fopen('fusion_rt.txt', 'a');
        fprintf (datafile,
'subname\tsubsex\tsubage\tnsub\tdate\ttime\ttrial\tinterval\tf0\trispSE\tri
spAC\trispRT\tISI\tOrder\n');
```

```
    else
        datafile = fopen('fusion_rating.txt', 'a');
    end
    % write all output values
    [rows, columns]=size(output);
    for i=1:rows
        fprintf(datafile, '%s\t%s\t%2.0f\t%2.0f\t%s\t%s\t', subname,
subsex, subage, nsub, date, time);
        for j=1:columns
            fprintf(datafile, '%4.1f\t', output(i, j));
        end
        fprintf(datafile, '%s\t', order);
        fprintf(datafile, '\n');
    end
    % close the file
    status = fclose (datafile);
end
```

Function used in the consonance rating task

```
function pref_rating(nsub, subname, subsex, subage, order)

%% psychtoobox settings
Screen('Preference', 'SkipSyncTests', 1);
Screen('Preference','SuppressAllWarnings');
screenNumber=max(Screen('Screens'));
black=BlackIndex(screenNumber);
white=WhiteIndex(screenNumber);
gray=round((white+black)/2);
% This makes sure that on floating point framebuffers we still get a
% well defined gray. It isn't strictly neccessary in this demo:
if gray==white
    gray=white/2;
end
rect1=Screen('Rect', screenNumber);
[w, rect]=Screen('OpenWindow',screenNumber, black, rect1);
%[w, ~]=Screen('OpenWindow',screenNumber, black, [0, 0, 400, 400]);
ifi=Screen('GetFlipInterval', w);
slack=ifi/2;
prioritylevel=MaxPriority(w);
rand('state', sum(100*clock));
Screen('TextFont', w, 'Lucida Sans');
Screen('TextSize', w, 24);
Screen('TextStyle', w, 1);
KbName('UnifyKeyNames');
oneKey=KbName('1!');
twoKey=KbName('2@');
threeKey=KbName('3#');
fourKey=KbName('4$');
% sound settings
```

```
sf = 48000;
d = 2000; % duration of the stimulus
InitializePsychSound;
pahandle = PsychPortAudio('Open', [], [], 0, sf, 1);

%% EXPERIMENT SETTINGS
ISI = [0.5,1];
% each stimulus is repeated 10 times
repetitions = 1;
    intervals = [203, 386, 498, 590, 702, 1088, 1200, 1312]; % size of the
intervals in cents
    f0 = [110, 110*2^(2/12), 220, 220*2^(2/12), 440, 440*2^(2/12)];

% here I create an event table that includes the characteristics of all the
% trials of the experiment
[event_table, ntrials]=makeeventtable(intervals, f0, ISI, repetitions);
trialorder=randperm(ntrials);
% trialorder=1:ntrials;

% the "output" variable is used later to store all the data of the
% experiment trial by trial
output=zeros(ntrials, 7);

%% experiment
WaitSecs(3);
HideCursor
Screen('Flip', w);
% main cicle of the experiment with all the trials
for trial=1:ntrials
    WaitSecs(1);
    if trial== ntrials/2
        DrawFormattedText(w, 'Premere la barra spaziatrice per continuare',
'center', 'center', white);
        Screen('Flip', w);
        while 1
            [~, ~, keyCode] = KbCheck;
            if keyCode(KbName('space'))
                break;
            end
        end
        WaitSecs(1);
    end
    % generate the string for the current trial
    s = SynthesizeSound(sf, d, event_table(trialorder(trial), 1),
event_table(trialorder(trial), 2));
    % show the fixation point
    DrawFormattedText(w, '+', 'center', 'center', white);
    PsychPortAudio('FillBuffer', pahandle, s');
    timeElapsed = 0;
    % show the trial at maximal priority
    Priority(prioritylevel);
    % keep the fixation for 180 ms
    t=Screen('Flip', w);
```

```
    t0=Screen('Flip', w, t+0.5-slack);
    PsychPortAudio('Start', pahandle, [], t0+event_table(trialorder(trial),
3));
    % collect response
    [~, rispRT, kbCode] = KbCheck;
    while kbCode(oneKey)==0 && kbCode(twoKey)==0 && kbCode(threeKey)==0 &&
kbCode(fourKey)==0
        [~, rispRT, kbCode] = KbCheck;
        timeElapsed = rispRT - t0-event_table(trialorder(trial), 3);
    end
    Priority(0);
    PsychPortAudio('Stop', pahandle);
    % store results
    output(trial, 1)=trial;
    output(trial, 2)=event_table(trialorder(trial), 1);% interval
    output(trial, 3)=event_table(trialorder(trial), 2);% f0
    if find(kbCode)
        responses = find(kbCode);
        output(trial, 4)=responses(1);
    else
        output(trial, 4)=999;
    end
    WaitSecs(3);

    if kbCode(oneKey)
        risp = 1;
    elseif kbCode(twoKey)
        risp = 2;
    elseif kbCode(threeKey)
        risp = 3;
    elseif kbCode(fourKey)
        risp = 4;
    else
        risp = 0;
    end

    output(trial, 5)= risp;
    output(trial, 6)= 1000*timeElapsed;
    output(trial, 7)= event_table(trialorder(trial), 3);
end
%% end of experiment
if nsub==0
    DrawFormattedText(w, 'La prova è finita', 'center', 'center', white);
else
    DrawFormattedText(w, 'L''esperimento è finito', 'center', 'center',
white);
end
Screen('Flip', w);
WaitSecs(3);
Screen('Flip', w);
ShowCursor;
Screen('CloseAll');
PsychPortAudio('Close', pahandle);
```

```
%% write datafile
% check whether the datafile exists in order to write if necessary the
% names of the variables
time=clock;
time=time(4:5);
time=[num2str(time(1)),':',num2str(time(2))];
if nsub ~= 0
    if ~exist('fusion_rt.txt', 'file')
        datafile = fopen('fusion_rt.txt', 'a');
        fprintf (datafile,
'subname\tsubsex\tsubage\tnsub\tdate\ttime\ttrial\tinterval\tf0\trispSE\tri
sp\trispRT\tISI\tOrder\n');
    else
        datafile = fopen('pref_rating.txt', 'a');
    end
    % write all output values
    [rows, columns]=size(output);
    for i=1:rows
        fprintf(datafile, '%s\t%s\t%2.0f\t%2.0f\t%s\t%s\t', subname,
subsex, subage, nsub, date, time);
        for j=1:columns
            fprintf(datafile, '%4.1f\t', output(i, j));
        end
        fprintf(datafile, '%s\t', order);
        fprintf(datafile, '\n');
    end
    % close the file
    status = fclose (datafile);
end
```

Function to generate the trials and stimuli combinations

```
function [event_table, number_of_trials] = makeeventtable(factor1, factor2,
factor3, factor4, factor5)

if nargin == 4
    factor5 = 1;
elseif nargin == 3
    factor5 = 1;
    factor4 = 1;
elseif nargin == 2
    factor5 = 1;
    factor4 = 1;
    factor3 = 1;
elseif nargin == 1
    factor5 = 1;
    factor4 = 1;
    factor3 = 1;
    factor2 = 1;
end
```

```
% number_of_trials = length(lag)*length(sound_pos)*length(repetitions);
number_of_trials =
length(factor1)*length(factor2)*length(factor3)*length(factor4)*length(fact
or5);
% I add 1 so that I can add also the randper of trials
event_table = zeros(number_of_trials, nargin + 1);

i = 1;
for f1 = 1:length(factor1)
    for f2 = 1:length(factor2)
        for f3 = 1:length(factor3)
            for f4 = 1:length(factor4)
                for f5 = 1:length(factor5)
                    event_table (i, 1) = factor1(f1);
                    event_table (i, 2) = factor2(f2);
                    event_table (i, 3) = factor3(f3);
                    event_table (i, 4) = factor4(f4);
                    event_table (i, 5) = factor5(f5);
                    i = i + 1;
                end
            end
        end
    end
end
```

Function to generate the sound stimuli

```
function s = SythesizeStimulus(sf, d, interval, f0)

esponente = [0, -14, -22.2, -28, -32.5, -36.2, -39.3, -42, -44.4, -46.5, -
48.4, -50.2];
amp = 10.^(esponente/20);                                                 %
amplitude of every sound harmonic
f_low = f0:f0:f0*12;
f_high = f0 * 2^(interval/1200);

% ramp
dBperduration = -24;                                                      %
generate the exponential decay -4s^-1(ramp)
dBperduration = log(10^(dBperduration/20));
ramp = exp(dBperduration*((1:1:d*sf/1000)/(d*sf/1000)));
ramp = (ramp / max(abs(ramp)))';

% synthesize sound
% lower tone
s1 = ramp .* GenerateTone(sf, d, f_low, amp);
s2 = ramp .* GenerateTone(sf, d, f_high, amp);
s = AttenuateSound(GenerateEnvelope(sf, (s1+s2)/max(abs(s1+s2))), -15);
```