UNIVERSITY OF PADOVA

DEPARTMENT OF GENERAL PSYCHOLOGY

Doctoral course in Psychological Sciences

Cycle: XXXIV

MUSIC AND COGNITION: THE EFFECT OF MUSIC LISTENING AND MUSIC TRAINING IN AGING

COORDINATOR: Ch.mo Prof. GIOVANNI GALFANO
SUPERVISOR: Ch.mo Prof. ROSSANA DE BENI
CO-SUPERVISOR: Ch.mo Prof. CHRISTIAN AGRILLO

Ph.D. STUDENT: Margherita Vincenzi

Index

Abstract

The relationship between music and cognition represents a particularly interesting issue in cognitive psychology field. In particular, research concerning music and cognition refers to two areas of study: the first one includes research on listening to music and cognitive performance (before or after the listening condition); the second one investigates the impact of music training (that is, have studied music) on cognition. Although in recent years there has been a gradual and progressive increase in studying the relationship between music and cognition in older age, little is known about whether and how music may modulate and influence cognitive and affective performance in the adult life span and in older adults. Moreover, as regards measuring musical abilities, considering the technological opportunities and the global spread of the internet, it becomes therefore particularly important to verify whether musical abilities can be successfully measured remotely online in a reliable way. Hence, the goal of the present thesis was to investigate the relationship between music and cognition across the adult life span, with particular attention to older people, considering on the one hand the effect of listening music on cognitive performance, and on the other the influence of music training on cognition. In order to deepen these topics, the three studies presented here aimed to: i) explore the effect of music listening on cognitive performance and on the emotional state of typical aging older adults; ii) investigate the relationship between music training, cognitive abilities and personality in professional musicians, iii) verify whether a musical skill test could be successfully administered online. Specifically, Study 1 examined the influence of music listening on the emotional state and cognitive tasks in healthy typically aging older adults. Study 2 aimed to investigate the association between musical abilities, cognitive performance and personality in professional, non-professional and non-musicians. Finally, Study 3 focused on the measurement of musical abilities using a test proposed online, compared to the same test administered in laboratory. The results showed that listening to music with particular musical characteristics can benefit the emotional state (i.e., mood, arousal), and cognition (i.e., visuo-spatial working memory) in healthy typically aging older adults. In addition, the "profile" of the professional musicians was further characterized at personality and cognitive level. Finally, the possibility of successfully measure musical abilities online was confirmed. In the thesis, the theoretical background, methodology, statistical analysis, results and general discussions are provided, along with general conclusions and implications.

Introduction

Music is part of our everyday lives: our environments constantly subject us to musical stimuli from the environment, from the moment of birth and across the entire life span. The relationship between music and cognitive performance has long been a focus of the field of cognitive psychology (Radocy & Boyle, 2012). In particular, the effect that music has on cognition is still a realm of active discussion, and the literature in this field diverges into two main lines of research: studies concerning *music listening* and cognitive performance, and studies focusing on *music training* and cognitive abilities. The first type of studies examines the effect that listening to music has on a person's cognitive performance, while the second investigates whether studying music has an impact on cognition.

Recent decades have seen an increase in population of older adults (especially in the more developed countries); therefore, exploration of the relationship between music and cognition is of particular interest, given music's potential as a protective factor in supporting the cognitive well-being of older adults as they age (Balbag et al., 2014). However, studies of these aspects in older age are still limited in number; considering the potential that lies in promoting older adults' cognitive enhancement through music listening in the older people (Laukka & Juslin, 2007; Silva et al., 2020), an understanding of how music might sustain older adults' cognitive performance is an increasingly important research aim. Moreover, it is of particular interest to further understand whether and to what extent the lifelong study of music may act as a protective factor during the gradual and continuous process of natural cognitive change that begins in adulthood and continues with age.

The present dissertation was designed to examine, with a particular focus on older adults, the relationship between music and cognition in the adult life span, considering on one hand the effect of listening to music on affective state and cognitive performance, and on the other hand the influence of music training on cognitive abilities across the adult life span.

In Chapter 1, several theoretical frameworks for music's effect cognition are considered, starting from the original so-called Mozart effect (Rauscher et al., 1993)—that is, an improvement in visuospatial tasks performance observed after listening to Mozart music—to the currently most accepted hypothesis to explain that effect, the arousal and mood hypothesis (Thompson et al., 2001) and the most recent findings in support thereof. After a further investigation of the literature on music listening and cognitive abilities in aging, this chapter reviews several studies of the relationship between music training and cognitive abilities, particularly focusing on older professional musicians and considering the role of personality traits in the music–cognition connection. Then, an overview of the existing measures to test musical abilities is proposed. Finally, it is discussed how to use the internet and online testing to measure musical abilities.

Chapter 2 describes the first study, which explored the effect of listening to music on affective states and cognition in a sample of healthy older adults. In Chapter 3 is presented the second study, which aimed to examine the relationship between music training, personality, and cognitive abilities across the adult life span. Chapter 4 illustrates the third and last study, which investigated the possibility of testing musical abilities online. Finally, Chapter 5 contains a general discussion of the relevant results and conclusions drawn on the findings of the studies conducted (see Table 1 for a brief summary of the content of the three studies). Appendixes 1 and 2 include the complete scores of the excerpts that were played for participants in Study 1.

All the three studies received the approval of the Ethical Committee for Psychological Research at the University of Padova, or of the ethics committee at ISCTE-IUL (Lisbon).

Table 1. *Overview of the content of the three studies.*

General aim(s)	**Sample**	**Materials & Procedure**	
		Affective state	*Cognitive tasks*
Study 1: Explore the effect of listening to music on older adults' affective state and cognitive tasks	132 healthy older adults (age range: 65-75 years)	Mood and arousal: SAM (Bradley & Lang, 1994)	VS-WM: backward Corsi Blocks Task (Corsi, 1972) EF: Verbal Fluency (Novelli et al., 1986); TMT-B (Amodio et al., 2002) Arithmetical abilities: AC-FL (Caviola et al., 2016).
		Objective behavioral tests	*Questionnaires*
Study 2: Investigate the relationship between music training and cognitive ability in the adult life span	642 adults (age range: 18-84 years)	Musical ability: MET (Wallentin et al., 2010) General cognitive ability: MaRs-IB (Chierchia et al., 2019)	Musical expertise: Gold-MSI (Müllensiefen et al., 2014) Personality: BFI (John & Srivastava, 1999) Mind Wandering: MWQ (Mrazek et al., 2013)

		Objective behavioral tests	*Questionnaires*
Study 3: Determine whether an objective test of musical ability can be successfully administered online	608 adults (age range: 18-88 years)	Musical ability: MET (Wallentin et al., 2010) General cognitive ability: MaRs-IB (Chierchia et al., 2019)	Musical expertise: Gold-MSI (Müllensiefen et al., 2014) Personality: BFI (John & Srivastava, 1999) Mind Wandering: MWQ (Mrazek et al., 2013)

Note. SAM: Self-Assessment Manikin; VS-WM: Visuospatial Working Memory; EF: Executive Functions; TMT-B: Trail Making Test-B; AC-FL: arithmetic fluency; MET: Musical Ear Test; MaRs-IB: Matrix Reasoning Item Bank; Gold-MSI: Gold Musical Sophistication Index; BFI: Big Five Inventory; MWQ: Mind Wandering Questionnaire.

1. Music and Cognition: theoretical framework and insights from the literature

1.1 Music and cognitive abilities: An introduction to the terms

Music can be defined as “the science or art of ordering tones or sounds in succession, in combination, and in temporal relationships to produce a composition having unity and continuity” (Merriam-Webster’s Collegiate Dictionary, 2022). As one of the universal cultural aspects pervading every human society, music accompanies an individual throughout the life span. The implication of music in a variety of fields (e.g., advertising, geriatrics, psychology) testifies to the extent to which music permeates human society and to its potential consequences for behavior (Rabinowitch, 2020) and cognition (Deliège & Sloboda, 1997).

Research about music and cognition could advance our understanding of the variety of processing resources and abilities involved in music listening. Indeed, the relationship between music and “cognitive abilities”—(i.e., all aspects of cognition, such as memory, visuospatial abilities, language, and general intelligence; Schellenberg & Weiss, 2013)—has been an issue of interest for the last 50 years, and it is still an area of active discussion. Literature in this area can be divided has diverged into (a) studies about listening to music and cognitive performance, and (b) studies focusing on musical training and cognitive abilities.

The first type of studies, over the years, evolved to discern the effect of listening to music on the person’s execution of cognitive tasks, and to examine whether and how it could enhance cognitive performance. Studies of the second type comprise a large body of research, examining the association between music lessons and non-musical cognitive abilities, which aims to determine whether musical training improves cognition. Most studies have correlational or quasi-experimental designs, precluding conclusions of causation (Schellenberg, 2019). Thus, it remains a matter of great debate whether and under what conditions musical training may improve nonmusical abilities.

Overall, research on the relationship between music and cognition is a particularly interesting field that could illuminate the variety of abilities involved in music processing.

Moreover, the potential influence of music on human behavior, emotions and cognition may play a crucial role in supporting individuals in everyday life.

The following two paragraphs will further discuss these two main lines of research, that is, the study of the music *listening*'s effect, and that of music *training*'s effect. It is, thus, of interest to consider the way in which music (i.e., music listening, and/or music training) relates to aging, which will be discussed more in deep in the successive sub-sections (1.2.1; 1.3.1).

1.2 From the original Mozart effect to the arousal and mood hypothesis and the most recent findings

The question of whether music listening can enhance cognitive performance is still debated. In particular, the effects of music listening on cognition have been studied through two different music presentation modalities: (a) with music played either *before* starting a task, and (b) with music played *while* completing a task (i.e., background music).

The original study that identified the famous, debated Mozart effect used the first music presentation modality (Rauscher et al., 1993): 36 students completed three visuospatial reasoning tasks after experiencing three different listening conditions (i.e., listening to Mozart's sonata K 448, relaxation instructions, or silence). Results showed that cognitive performance across tasks was significantly higher in the Mozart-listening condition than in the other two conditions, and this finding became known as the Mozart effect. This effect had an immediate impact on the research community, leading to the idea that listening to Mozart could make a person smarter. However, fewer than half of the studies trying to replicate Rauscher's findings were able to find the same pattern produced by the original study (see Chabris, 1999; Hetland, 2000, and Pietschnig et al., 2010 for meta-analyses). Moreover, the effect seems to be limited in time: usually, it can be observed within 10 to 15 minutes after listening (Jenkins, 2001), and has been observed both with musical stimuli (e.g., Schellenberg

& Hallam, 2005; Schellenberg et al., 2007), and nonmusical stimuli (for instance, audiobooks; see Nantais & Schellenberg, 1999).

Thompson et al. (2001) advanced a possible explanation of the Mozart effect within the framework of the arousal and mood hypothesis, which suggested that any stimulus can affect how a person feels and, subsequently, cognitive performance. Specifically, the arousal and mood hypothesis postulates that the effect of music listening on cognitive performance, rather than being the product of listening to a given piece (e.g., Mozart's sonata K 448), could be related more broadly to the impact of music on the emotional state of the listener, which could in turn influence cognitive performance. This idea is also in accordance with evidence of the relationship between music listening and emotional response: listening to music changes the way a person feels (e.g., Juslin et al., 2008; Sloboda, 1992) by, for example, lowering cortisol level (Flaten, et al., 2006) or blood pressure (Triller et al., 2006) in the case of calming music, or by facilitating falling asleep (Dickson & Schubert, 2019).

According to the arousal and mood hypothesis, the specific harmonic characteristics of a music piece, such as tempo (fast or slow) and mode (major or minor), modulate the listener's arousal (i.e., the degree of physiological activation) and the mood (i.e., persistence of emotions), respectively, with a positive effect on the individual's emotional state and on the subsequent cognitive performance (Husain et al., 2002; Schellenberg et al., 2007).

Mood-induction literature (e.g., Västfjäll, 2001) seems to have confirmed the influence of listening to music on self-reported mood. Music can trigger deep emotions in the listener (Juslin & Sloboda, 2001), and emotional recognition in music is a common and almost "automatic" process (Peretz et al., 1998), observable from childhood (Dalla Bella et al., 2001; Nieminen et al., 2011); however, listeners also expressed emotions after hearing an unfamiliar specific musical system (Fritz et al., 2009). Therefore, mood-induction research seems to be in accordance with the idea that some structural music properties—such as the abovementioned tempo and mode—contribute to eliciting certain emotions in the listener (for reviews see Gabrielsson & Lindström, 2001; Juslin et al., 2008).

The second music presentation modality, background music, exposes listeners to music while they complete tasks. Research on background music mainly focuses on whether the listening affects performance on a concurrent primary task; the mixed results in this field have found background music to have beneficial, detrimental, or no effect on a variety of psychological outcomes (see Kämpfe et al., 2011, for a meta-analysis). However, a comparison of different types of background music revealed that the tempo of the music (i.e., fast or slow) influenced participants' performance on the tasks they executed while being exposed to music (e.g., Bottiroli et al., 2014). In addition, the large individual differences (e.g., personality) and contextual factors (e.g., task difficulty) make results equivocal, and background music has sometimes appeared to enhance performance in certain areas (e.g., in sports achievements and affect), whereas it has had detrimental effects on other aspects (e.g., on reading comprehension and memory; Kämpfe et al., 2011).

Given the evidence of music's crucial role in supporting emotions (Juslin, 2019), and thereby affecting a fundamental part of individuals' everyday functioning, it is also of interest to consider the potential positive effect of music listening in everyday life. Moreover, considering that Western society is getting considerably older, and that the age-related decline of certain cognitive abilities poses a threat to the autonomous well-being of the elderly, the identification of a means to sustain and promote successful aging is fundamental. The effect of music listening on older adults will be discussed in the next paragraph.

1.2.1 Music listening and cognitive abilities in aging: what does the literature suggest?

As mentioned in previous paragraphs, the literature on cognitive abilities and listening to music has demonstrated music's ability to modulate listeners' emotional states and impact on their cognitive performance. This ability of music to influence and trigger emotional phenomena seems to last throughout the life span, from childhood to old age. Over the last few decades, studies on aging have focused on proposing interventions to defer age-related cognitive decline and promote healthy aging. Considering the rise of the population's average age, a clear picture of whether and how music may sustain older adults' cognitive and emotional well-being represents a very important issue.

However, it must be said that studies of the effect of listening to music on older adults' cognitive performance still lack stability on both empirical and theoretical grounds, and that the data in support of those studies' claims are sparse (e.g., Silva et al., 2020). Further, few studies address this issue in older adults and they present contrasting results. Specifically, studies have found music listening to have positive and negative effects on verbal fluency (Thompson et al., 2005, Giannouli et al., 2019, respectively); a positive effect on source memory (Reaves et al., 2016; Palumbo et al., 2018); positive (Wang et al., 2013; Mammarella et al., 2007), negative (Hirokawa 2004; Giannouli et al., 2019), and null effects (Borella et al., 2014; Borella et al., 2019) on working memory (WM) and WM training; and a positive effect on memory encoding (Ferreri et al., 2014), episodic memory and processing speed (Bottiroli et al., 2014), spatial-temporal abilities (Cacciafesta et al., 2010; Padulo et al., 2020), and negative-affect regulation (Groarke & Hogan, 2019).

In this varied picture, it is worth mentioning that, in the literature, a consensus has emerged that emotional response to music is intact in older adults. For example, after negative affect is induced, self-selected music has a stronger ameliorating effect for older adults than for their younger counterparts (Groarke & Hogan, 2019).

However, one reason of these contrasting results (and a likely contributor to inconsistencies across studies), could be related to studies' large heterogeneity in terms of pieces, genres, and durations of the selected music excerpts (see Table 2).

Some studies used two (or more) different music pieces (e.g., Bottiroli et al., 2014); in others, participants only listened to one music excerpt (e.g., Mammarella et al., 2007). In some studies, researchers asked participants to bring their own music (e.g., Groarke & Hogan, 2019), but did not consistently provide instructions on the characteristics the music pieces should present. The use of several music pieces, regardless of their mode or tempo (which, as introduced previously, exert a crucial element of influence on participants' cognitive performance through mood and arousal changes), may have contributed to the confusion in the studies' results.

The genre from which studies drew their music excerpts ranged from Baroque music (e.g., Mammarella et al., 2007: Vivaldi's *Four Seasons – Spring*) to jazz (e.g., Ferreri et al., 2014: Bechet's 'If You See my Mother').

Finally, the duration (length) of the listening condition proposed in the studies was extremely diverse, varying from a few seconds (e.g., Ferreri et al., 2014), to 10 minutes listening (e.g., Hirokawa, 2004). In addition, as a further issue, literature suggests that the duration of the effect of music listening lasts for the 12 minutes after the listening condition (Jenkins, 2001), and then gradually disappears; that this aspect has also rarely been considered in previous studies contributes to the mixed findings.

In addition, the two abovementioned music presentation modalities, that is, before the cognitive tasks (as in mood-induction literature), and in the background of cognitive tasks, eventually merged in studies of older adults, and the different effect that the presentation modality has on the listening experience has rarely been considered.

Although the findings of studies on listening to music and cognitive abilities in the older adults are mixed, the arousal and mood hypothesis—at present, the most plausible explanation of the effect of music listening on cognition—has not yet been directly tested on older adults.

Table 2. *Summary of the studies regarding music listening and cognitive performance including older adults.*

Author, year	N	Age	Measures	Experimental design	Control group characteristics	Listening condition	Listening condition design	Listening condition duration	Outcomes
Hirokawa, 2004	15 OA	M = 72,7 y	**Arousal:** Activation-Deactivation Adjective Check List **WM**: reading span test	Between subjects' design	1 Relaxation instructions 1 No music (silence)	Participants preferred instrumental music recordings or provided by the experimenter	**Prior to the task**	**10 minutes**	Subject-preferred music: + arousal + relaxation, – tension Relaxation: – tension Silence: – energy levels No differences in WM No relationship arousal levels / WM scores
Thompson et al., 2005	32: 16 OA 16 Alzheimer's (AD) OA	M_{OA} = 74,94 y M_{AD} = 76,25 y	**EF:** Category verbal fluency	Repeated-measure mixed design	No music (silence)	Vivaldi's Winter – Four Seasons	**Background**	**≈ 1 min and 10 s**	Music: + category verbal fluency in healthy older adults and AD
Mammarella et al., 2007	24	M = 81 y	**Short term memory**: Forward digit span **EF:** Verbal fluency	Within subjects' design	White noise	Vivaldi's Spring – Four Seasons	**Background**	**≈ 8 minutes**	Music condition > white noise condition: + EF

Author, year	N	Age	Measures	Experimental design	Control group characteristics	Listening condition	Listening condition design	Listening condition duration	Outcomes
Cacciafesta, 2010	12 MCI	Age range 66-77 y	**Spatial abilities**: PFC **Episodic learning**: 3 objects and 3 places test **Ideational-praxis abilities**: clock-drawing test **Immediate recall**: Rey's 15-word test **Mental flexibility and speed**: TMT (A-B) **Short term / WM**: Forward and backward digit span	Within subjects' repeated measures design	Same patients enrolled as control group	- Mozart K 448 - Beethoven Fur Elise	**Prior to the task**	**≈ 9 minutes**	Music listening in MCI: + spatial–temporal abilities + immediate recall
Ferreri et al., 2014	16	M = 64,5 y	**Source memory**: list of words	Within subjects' design	No control group, silence condition	"If you see my mother", (S. Bechet)	**Background**	**≈ 8 minutes**	Music condition: + source-memory – decreased PFC activity
Bottiroli et al., 2014	65	M = 69,03 y	**Episodic memory**: list of words **EF**: phonemic verbal fluency **Mood questionnaire**	Within subjects' design	1 No music (silence), 1 White noise	- Mozart Eine-kleine Nachtmusik, - Mahler's Adagietto	**Background**	**≈ 10 minutes**	Music: > silence and > white noise: + memory performance Mozart group: + processing speed
Borella et al., 2014	63 YA 92 OA	M_{YA} = 25,60 y M_{OA} = 68,60 y	**Short term / WM**: CWMS, forward and backward Digit Span **Emotions / Working memory**: Affective Ospan **Processing speed**: Pattern comparison **Mood**: PANAS	Between subjects' design	Neutral short description (television invention)	- Mozart Sonata K 448 - Albinoni Adagio in G minor	**Prior to the task**	**≈ 10 minutes**	Music: – young adults' performance Albinoni condition: – young adults' performance
Reaves et al., 2016	53 YA 50 OA	M_{YA} = 20.9 y M_{OA} = 66.3 y	**Memory Assessment Scale battery:** forward and backward digit span, verbal list learning, face–name paired recognition task **Speed and attention:** TMT-A, -B **EF:** Verbal fluency, Controlled Oral Word Association test ("FAS") **Spatial WM span:** Corsi blocks task	Between subjects' design	1 No music (silence) 1 "Musical rain" (computer generated vowel sound)	Music stimuli: selected from experimenters' personal collections and online (182 songs)	**Background**	**15 seconds**	Music: – older adults' associative memory

Author, year	N	Age	Measures	Experimental design	Control group characteristics	Listening condition	Listening condition design	Listening condition duration	Outcomes
Borella et al., 2019	72	M = 69,24 y	**Mood:** qualitative interview **EF:** verbal fluency **WM**: CWMS **Vs-WM**: backward Corsi blocks **Spatial visualization**: MPFB **Spatial learning**: spatial descriptions **Fluid intelligence**: Cattell test	Between subjects' design	1 white noise 1 No music (silence)	- Mozart Sonata K 448; - Albinoni Adagio	**Prior to the task**	**6 minutes**	No music effect on WM training Albinoni music: larger specific training gains in the criterion task
Palumbo et al., 2018	144 YA 144 OA	M = 22,2 y M = 68,3 y	**Short term / WM:** Forward and backward digit span **EF:** verbal fluency **Affective state**: PANAS **Emotions**: IAPS pictures	Between subjects' design	White noise	- Mozart Eine-kleine Nachtmusik - Mahler's Adagietto	**Background**	**5,30 minutes**	Music: + in both younger and older adults Classical-music condition > white noise condition: + source memory
Giannouli et al., 2019	240 YA 227 OA	M_{YA} = 28,63 y M_{OA} = 72,23 y	**EF:** verbal fluency **Short term memory**: forward digit span	Within subjects' design In groups of 10 participants	No music (silence)	- Mozart Sonata K 448, - Vivaldi harpsicord op.4 n 10, - Philip Glass Music with changing parts	**Prior to the task**	**10 minutes**	Vivaldi: – verbal WM in both groups + verbal fluency in young adults Mozart: - verbal fluency in both groups
Groarke et al, 2019	40 YA 40 OA	M_{YA} = 19.75 y M_{OA} = 68.48 y	**Subjective affect:** VAS **Stress:** TSST (+ mental arithmetic task)	Between subjects' design	Radio documentary	- Self-selected music for stressful situations	**Prior to the task**	**10 minutes**	Music listening condition > control condition + affect regulation score
Padulo et al., 2020	179 YA, 183 OA	M_{YA} = 22.51 y M_{OA} = 74.15 y	**Short term memory**: forward digit span **WM**: backward digit span, **EF**: verbal fluency test (FAS) **Mood**: PANAS, SAM **Spatial visualization and motor skil**l: block design subscale (WAIS-R)	Between subjects' design	White noise	- Mozart Sonata K 448 (until beat n. 80)	**Prior to the task**	**5 minutes**	Music: – no effect on young adults' visuo-spatial tasks + accuracy and mean times after Mozart listening White noise: – spatial reasoning in older adults

Author, year	**N**	**Age**	**Measures**	**Experimental design**	**Control group characteristics**	**Listening condition**	**Listening condition design**	**Listening condition duration**	**Outcomes**
Silva et al., 2020	12 OA	M = 75.25 y	**Long-term memory:** word list **Processing speed:** Letter comparison task	Within subjects' design	1 Silence 1 Environmental sounds	- Music 1/metal ("John and the Creatures – Here's to the Crazy Ones" - Music 2/electronic "Robert Miles – Children" - Music 3/jazz "Thelonious Monk - Blue Monk"	**Prior to the task**	**20 seconds**	Music: – no effect in young adults + memory in older adults
Chow et a., 2021	14 OA	M = 72.6 y	**Affective state**: PANAS **Short term /WM:** Forward and backward digit span **Recognition memory**: Auditory word recognition task **tDCS stimulation** (tDCS + music / tDCS only / Sham + Music)	Within subject's repeated measures design	No music (silence)	Vocal / instrumental music selected by participant and experimenter (classical, rock, jazz, folk, pop, country, and film score)	**Background**	**21 minutes**	Backwards digit span: + in tDCS + Music, / Sham + Music conditions. No differences in auditory WRT, Recognition memory: after tDCS + Music
Ward et al., 2021	48 YA 48 OA	M_{YA} = 19.65 y M_{OA} = 73.92 y	**Memory encoding:** free recall of 20 words **WM**: backward digit span **Mood induction:** Geneva Affective Picture Database	Within subject's repeated measures design	No control group	Originally composed music	**Prior to the task**	**2 min**	Recall: + after mood-matching in young and older adults, – after mood-mismatching music in older adults WM: + in mood-matching condition, no differences from baseline in mismatching condition

Note. YA: Older Adults, OA: Older Adults. MCI: Mild Cognitive Impairment. AD: Alzheimer Disease. MMSE (Mini Mental State Examination); NART (National Adult Reading Test); PFC (Paper-Folding and Cutting test); TMT (Trail Making Test); CES-D (Center for Epidemiological Studies Depression) ; SVAMA (Italian Checklist for the Multidimensional Assessment); CWMS (Categorization working memory span task); PANAS (Positive and Negative Affective Scale); CDR (Clinical Dementia Rating scale); Minnesota Paper Form Board (MPFB); GDS (Geriatric Depression Scale); AFML (Adaptive Functions of Music Listening); TSST (Trier Social Stress Test); VAS (Visual Analogue Scale); SAM (Self-Assessment Manikin); WAIS (Wechsler Adult Intelligence Scale)

1.3 Literature and open issues regarding music training and cognitive abilities

Another line of research in the music and cognition field has focused on the effect of musical training on nonmusical cognitive abilities. Specifically, such research addresses whether and how professional musical training has a positive effect on cognitive abilities that are not necessarily music-related. In fact, a considerable number of studies have tried to discern whether musical training improves nonmusical cognitive abilities, as musical training has positive associations with visuospatial, language, and general cognitive abilities tasks' (see Swaminathan & Schellenberg, 2019, for a review).

The literature suggests that musically trained participants outperform their untrained counterparts on a variety of tests of musical cognition. For instance, young and older adults with musical training seem to perform better than untrained participants do at recognizing melodies presented in transportation (Halpern et al., 1995) or at an abnormally fast or slow tempo (Andrews et al., 1998), and in detecting unusual notes in a familiar melody (Schellenberg, 2011).

Moreover, musically trained participants seem to be also better at detecting changes in pitch to the final word of a sentence and in a melody (Besson et al., 2007), at perceiving speech-in-noise (Parbery-Clark et al., 2009), and at using pitch patterns to identify words (Wong & Perrachione, 2007). Hence, musically trained participants seem to be very good listeners (Kraus & Chandrasekaran, 2010; Strait et al., 2010). However, since most of the studies in this field are quasi-experimental or correlational—which precludes inferences of causation—and although it is reasonable to assume that performing music might improve a person's listening ability, the reverse causal direction is equally plausible (Schellenberg et al., 2013, 2019). High-functioning individuals are more likely than others to begin musical training early and to take music lessons for many years. Indeed, individuals with lower listening abilities are less likely to start and/or continue musical training, which requires years of practice particularly in listening tasks specifically; those with high listening abilities might find easier to perform music tasks and, therefore, to persevere in musical training.

Therefore, given that it is impossible to infer causation starting from correlations, this issue is particularly difficult to solve.

Another aspect that makes the literature on this topic particularly complicated is that the criteria for defining a professional musician, which have only recently been deeply discussed (see Zhang et al., 2020). Currently, a consensus has settled on the “six-year rule” (i.e., at least 6 years of musical expertise) as a threshold for musical expertise (Zhang et al., 2020). However, it is worth mentioning that this rule ignores whether individuals are working or have worked as musicians, failing in distinguishing professional musicians, whose daily behaviors are dedicated to music, from musically trained individuals who are employed in other fields.

Having worked as professional musician in one’s lifetime also seems to have positive effects in aging. Discussion on this topic will be deepened in the next section.

1.3.1 Older professional musicians: does music training have benefits in aging?

Though age-related changes in cognitive abilities are well documented (e.g., Borella et al., 2008, Park et al., 2003), little is known about whether practicing music during one’s lifetime benefits on cognition in old age. Some evidence showed that older musicians achieve higher scores than nonmusicians in verbal memory (immediate recall) and executive processes (Hanna-Pladdy & Gajewski 2012; Hanna-Pladdy & MacKay, 2011). Older musicians seem to show more efficient cognitive control over irrelevant information than nonmusicians do (Amer et al., 2013). However, there seems to be no difference between musicians and nonmusicians’ visuospatial memory (Hanna-Pladdy et al., 2012); the cognitive profile of older musicians has yet to be further investigated.

It is worth mentioning that—as previously noted—studies considering older musicians suffer a lack of definition of musical expertise: some researchers considered musicians to be persons with 10 or more years of experience playing music, whereas others have looked to the age of beginning musical training (e.g., before 9 years old), or having played a musical instrument throughout life to define musical expertise for the purposes of their studies.

For example, Grassi et al. (2017) found that, with respect to nonmusicians, musicians (in that study, persons with between 46 and 80 years of musical practice, and who were still actively playing) had better cognitive (WM, short-term memory, and visuospatial abilities) and auditory (absolute threshold, frequency intensity, duration and spectral shape discrimination, gap and sinusoidal amplitude-modulation detection) profiles. Similarly, Hanna-Pladdy et al. (2011) reported better nonverbal memory and executive processes in older musicians (i.e., individuals who had at least 10 years of musical experience) compared to nonmusicians. Hence, older musicians seem to outperform nonmusicians on mainly near-transfer tasks (e.g., speech-in-noise perception and auditory WM; see Parbery-Clark et al., 2011), whereas evidence of far transfer (e.g., cognitive flexibility and nonverbal memory; Hanna-Pladdy et al., 2011) was less consistent.

Early musical training seems to be a protective factor against cognitive decline, due to the cognitive stimulation inherent in musical activity. Indeed, cognitive reserve—a resource to be used when the brain experiences increased burden (Stern et al., 2003)—is one potential construct to support musical training as protective of later-life cognitive function. Because the studies including older musicians are a limited number and the criteria defining a professional musician are not consistent, their results must be considered carefully. In addition, it needs to be followed the same criterion to define a professional musician, in order to compare results and performance.

1.3.2 The role of personality traits in professional musicians

As noted previously, the studies on musicians lack homogeneity concerning the definition of a *professional* musician: the majority of the studies only take into account the number of years of practice, leaving aside other fundamental factors, such as making music one's main job, along with demographic and personality characteristics. In fact, these latter two have found to be associated with musical training (Corrigall et al., 2013): it could be that individuals with a higher socioeconomic status are more likely to begin musical training early and to take music lessons for many years. Moreover, evidence suggests that personality traits predict occupational choices (Holland, 1997).

As for personality, one of the most dominant theoretical frameworks is the five-factor model, that is, the Big Five model (Goldberg, 1993; McCrae et al., 1999). In this, personality traits are conceptualized as individual's consistent pattern of thoughts, feeling and actions (McCrae, 1987), described by five major traits: emotional stability, extraversion, openness to experience, agreeableness, and conscientiousness. These traits are generally considered to be quite stable characteristics of an individual and to be relatively universal across different cultures (e.g., McCrae & Costa, 1997). Specifically, the openness to experience trait seems to be positively associated with creativity across domains (Feist, 1998, 2019; George & Zhou, 2001; Karwowski et al, 2016; Puryear et al., 2017). Creative activities seem to involve emotional aspects and cognitive processes, engaging the human brain in several ways: for example, the prefrontal cortex, involved in higher-order executive functioning (Friedman & Robbins, 2022), is also involved in divergent thinking, that is, a thought process or method used to generate creative ideas by exploring many possible solutions, representing a fundamental part of creativity (Heilman et al., 2003; McCrae, 1987). Moreover, various regions of an individual's parietal and temporal lobes (which have reciprocal connections with the frontal lobes) are bilaterally activated when the individual performs music (Gjermunds et al., 2020), implying that sensory information from various modalities (i.e., visual, auditory and proprioceptive information) can be processed and integrated with ongoing creative processes (López-González & Limb, 2012). Openness to experience also seems to predict musical behaviors and skills (e.g., Corrigall, et al., 2013; Lima et al., 2020) and time dedicated to practicing music (Butkovic et al., 2015).

Another personality trait that seems to predict creativity—though not as strongly as openness seems to (Feist, 2019)—is extraversion. For instance, Shuter-Dyson (2000) found that male and female music students scored higher on extraversion than male and female non-musicians did. Indeed, differences in personality traits such as extraversion have been associated with either musical genre and/or the person's gender (Rose et al., 2019): the social dynamics of musical learning may in fact have supported stereotypical patterns of findings in previous evidence (e.g., Cribb & Gregory, 1999).

However, it is worth noting that the majority of evidence gathered form investigations of personality traits in musicians comes from studies of rather small samples of musicians or music students of various musical genres (e.g., Butkovic & Modrusan, 2019). Moreover, the research designs of these studies differ significantly from one another, using a variety of personality tests: some studies used the16 PF (Buttsworth & Smith, 1995), others the EPQ-R (Shuter-Dyson, 2000), and others still the Myers-Briggs (MacLellan, 2011) or BFI-20 (Vaag et al., 2018). In addition, the referenced studies involve mainly young adults; a life span perspective that also includes older participants is still lacking.

In short, the studies examining personality in professional musicians (especially those including a broad age range) are still few. Therefore, the topic of personality in adult and older adult professional musicians merits further investigation.

1.4 An overview of the existing measures to test musical abilities

Because musical ability seems to be related to other nonmusical abilities (Radocy & Boyle, 2012), the question regarding how to measure it objectively has arisen over time. Indeed, the concept of musical ability —intended as musical predisposition, or musical *aptitude*— is a still debated matter, and determining the roles of nature and nurture remains an open issue (for a review, see Schellenberg, 2020). Considering that musical ability is now generally accepted as having both genetic and environmental influence, the term aptitude has mainly fallen out of use, in favor of more neutral terms (e.g., ability or competence). However, creating and providing accurate and objective measurement of musical ability becomes fundamental to trying to elucidate this topic, together with answering questions such as whether a single musical ability exist or whether it consists of a number of interrelated factors, and which those factors are.

Several researchers around the middle of the last century tried to develop musical ability batteries. The first attempts to measure musical ability were proposed in the early 1900s, in the frame of testing scholars' attitudes toward talent and nativism, but the subtests often measured a combination of skills rather than a specific one (Law & Zentner, 2012). A review of the most important measures of music ability is proposed in Table 3.

Seashore (1919) proposed one of the first tests of musical ability (Seashore et al., 1960), based on the idea that several perceptual subtests tapped different aspects of musical ability, and with the aim to determine whether a person was a suitable candidate for music training. Seashore claimed that each subtest of the Seashore Measures of Musical Talents measured one aspect of musical ability individually, without the possibility of being combined into one individual musical aptitude score. The Kwalwasser-Dykema Music Tests (1930) followed this same line. Wing (1948) then developed the Wing's Test of Musical Intelligence, consisting of seven perceptual subtests, whose scores were combined to yield an overall score of musical ability. In the Musical Aptitude Profile, Gordon (1965) sustained the concept of "developmental" musical ability, in which the level of musical aptitude a child is born with cannot be raised; after the age of 9, musical aptitude stabilizes (Gordon, 1979). Then, the notion of talent was questioned (Ericsson et al., 1993; Howe et al., 1998), and individual differences in musical achievement began to be considered as consequences of practice together with other environmental aspects (e.g., parental support).

After, a consistent number of listening tests have been developed, with a focus on academic research and suitability for an adult population. For example, Gordon subsequently simplified his approach in his Measures of Music Audiation, which were available in primary (kindergarten to Grades 3; Gordon, 1979), Intermediate (Grades 1– 6; Gordon, 1982), and Advanced (Grades 7–adult; Gordon, 1989) versions, providing separate scores for melody and rhythm. Gordon (1979) coined the term *audiation* to describe the process of retaining and comparing two musical sequences—a standard followed by a comparison—presented in succession. Since 2010, several measures of music abilities and expertise have been proposed, such as the Musical Ear Test (Wallentin et al., 2010), the Profile of Music Perception Skills (Law & Zentner, 2012), the Swedish Musical Discrimination Test (Ullén, et al., 2014), and the Harvard Beat Assessment Test (Fujii & Schlaug, 2013), considering music ability as a "consequence" of both genetic and environmental influences. Additionally, the Montreal

Battery for the Evaluation of Amusia (Peretz, et al., 2003) was developed with the aim of diagnosing congenital amusia (i.e., when musical abilities are congenitally low). Then, the Jake Mandell Tone Deaf Test (JMT, Palomar-García et al., 2019) and the Mandell Musical Hearing Tests were designed respectively to evaluate tone-deafness (congenital amusia) and to verify pitch perception and melodic memory, as well as identify neuroanatomical correlates of tone deafness.

Nowadays, most audio sample sounds used in previous tests, particularly those recorded in the first half of the past century, sound impure or distorted to the contemporary ear, due to limitations in recording techniques of the time or to the quality of the audio material having degraded over time (Law & Zentner, 2012). Another problem regarding previous measures concerns the overall design of the batteries, with an unequal number or duration of stimuli within a subtest, or variations in the answer format across subtests. Moreover, test–retest reliability was examined only occasionally. Therefore, some of the abovementioned measures are no longer used for research.

It is worth noting that the majority of these measures propose artificially created auditory experimental stimuli and rely on core musical abilities, specifically auditory short-term and/or auditory working memory and perceptual discrimination, without including a self-report music ability inventory. Therefore, the risk of encountering problems with ecological validity and low similarities with real music is particularly evident. In this framework, the Gold-Musical Sophistication Index (Gold-MSI) was designed in the context of "musical sophistication" (i.e., "A psychometric construct that can refer to musical skills, expertise, achievements, and related behaviours across a range of facets that are measured on different subscales"; Müllensiefen et al., 2014) to describe the multifaceted nature of musical expertise. Differently from the above-mentioned musical ability measures, which focus more on auditory discrimination, the Gold-MSI also considers other aspects of musical ability and behavior, providing information regarding active engagement with music, emotional responses to music, and self-reports of singing and perceptual abilities.

Therefore, multiple measures to test musical abilities are available, and it is necessary to choose carefully which of them to use during study design (considering their duration, and the abilities they test). In addition, putting together a subjective self-reported evaluation and an objective one measuring auditory abilities is advisable to capture both aspects.

Table 3. *Summary of music abilities measures.*

Author, year	Name of the measure	Goal	Subtests	Duration
Seashore, 1919	*Seashore Measures of Musical Talents*	Determine whether an individual is a suitable candidate for music training; aimed at practicing musicians	Pitch, Loudness, Time, Timbre, Frequency, Intensity, Duration, wave form	NA
Kwalwasser & Dykema, 1930	*Kwalwasser-Dykema Music Tests*	Examine music abilities in college student	Tonal Memory, Quality Discrimination, Intensity Discrimination, Tonal Movement, Time Discrimination, Rhythm Discrimination, Pitch Discrimination, Melodic Taste, Pitch Imagery, and Rhythm Imagery	NA
Wing, 1948	*Wing's Test of Musical Intelligence*	Musical ability, appreciation, tests acceptable to musicians. Production of musically meaningful responses to stimuli	Memory for pitch Rhythm	45-65 minutes
Drake, 1954	*Drake Music Aptitude Tests*	Relationship between years of musical training and performance	Musical Memory Rhythm	40 minutes
Gastón, 1957	*Test of Musicality*	Determine subject's ability to hear whether a given pitch is present in a chord, to detect differences between a heard versus a printed melody, to tell whether the final note of an incomplete melody should be higher or lower than the last note heard, and to note any pitch or rhythmic changes in a melody	Memory for pitch Rhythm	40/45 minutes
Bentley, 1966	*Measures of Musical Ability*	Investigate musicality (music acuity, musical hearing, and sensitivity to performance) in children	Pitch discrimination Tonal memory and Rhythmic Memory test Chord Analysis test	20 minutes
Gordon, 1965,1969	*Musical Aptitude Profile (MAP)*	Preference and non-preference subtests constitute the test battery and because, while the content of the test items is musical and is performed by professional musicians, the battery does, nevertheless, provide for the evaluation of seven postulated dimensions of musical aptitude	seven components, divided in three parts: tonal imagery (melody and harmony), rhythm imagery (tempo and meter), and musical sensitivity (phrasing, balance, and style).	110 minutes
Gordon, 1979 Gordon, 1982	*Primary Measures of Music Audiation (PMMA), Intermediate Measures of Music Audiatio (IMMA), and*	Music aptitude test for children and adults, for assessing musical talent (or aptitude)	Tonal part Rhythm part	25 minutes

Gordon, 1989	*Advanced Meansures of Music Audiation (AMMA)*			
Karma, 2007	*Karma Music Test*	Auditory structuring ability test; minimizes the effects of training and/or culture	Changes in patterns Changes in the order or number of the tones	NA
Wallentin et al., 2010	*Musical Ear Test (MET)*	Measuring musical abilities in both musicians and non-musicians	Melody subtest Rhythm subtest	20 minutes
Law & Zentner, 2012	*Profile of Music Perception Skills (PROMS)*	Individuals with musical skill, as well as those who, despite extensive musical training, may not be as skilled. It has also online format.	Tonal (melody, pitch), qualitative (timbre, tuning), temporal (rhythm, rhythm-to-melody, accent, tempo), and dynamic (loudness).	60 minutes
Ullén et al., 2014	*Swedish Musical Discrimination Test (SMDT)*	Provide measures of basic aspects of musical ability operationalized as discrimination ability for auditory musical stimuli	Discrimination of melodies, rhythms, and single pitches (similar to Bentley, MET, PROMS)	15 minutes
Fujii & Schlaug, 2013	*Harvard Beat Assessment Test (H-BAT)*	Battery of tests to assess beat perception and production abilities. Measures perception and production thresholds from the same set of auditory stimuli. Identify individuals who are performing below the cut-off scores and could be identified as beat-deaf	music tapping test (MTT), beat saliency test (BST), beat interval test (BIT), and beat finding and interval test (BFIT). The BST, BIT, and BFIT have a perception (per) and production (pro) part each.	35 minutes
Peretz et al., 2003	*Montreal Battery for the Evaluation of Amusia (MBEA)*	Diagnosing congenital amusia: battery of tests to screen perceptual problems in pitch, rhythm, and meter	contour, interval, scale, rhythm, meter, and memory tests	90 minutes
Mandell, 2006	*Jake Mandell Tone Deaf Test (JMT)*	Evaluate tone-deafness (congenital amusia), and to be challenging for subjects with musical training	Pitch discrimination	6 minutes

1.5 Online testing: how to test musical abilities using interactive Internet platforms

The internet has significantly changed the way people communicate and live, influencing the practice of psychology as it relates to testing and assessment (Naglieri et al., 2004). In particular, internet testing, compared to paper-and-pencil testing, lowered the cost of collecting data. Moreover, online testing is convenient, and it allows rapid communication of findings to clients, patients, researchers, and the public. However, a distinction between testing and psychological assessment (Matarazzo, 1990) has to be made because it becomes very important in this field: *testing* refers to the administration, scoring, and interpretation of individual test scores by applying a descriptive meaning based on normative, nomothetic data, whereas in *psychological assessment*, the emphasis is on the person being assessed, rather than on specific test results. This distinction is important because the majority of the content that is available online is testing, not psychological assessment. Therefore, in testing, the focus is on the individual test itself, and the aim is not to combine the result with a battery of other tests.

Indeed, online methods have been increasingly used as an alternative to in-person laboratory research (e.g., Chetverikov & Upravitelev, 2015; Houben & Wiers, 2008; Milne et al., 2020; Taherbhai et al., 2012), and several online platforms providing new tools for recruitment and testing have been created (e.g., Gosling & Mason, 2015; Grootswagers, 2020). In particular, the Covid-19 pandemic situation increased the possibility to reach people using the internet, restricting in-person contact and making online testing an interesting and attractive option for psychological research.

Although some concerns and possible problems exist, such as directly controlling the testing contexts, online testing shows some aspects that make it equivalent or even superior to in-person testing (e.g., Casler et al., 2013; Dandurand et al., 2008; Gosling et al., 2004). First, the internet allows researchers to reach a sample that is more diverse and representative in terms of age, gender and socioeconomic status; in addition, the access to relatively rare targets (e.g., musicians) is easier. Second, participants may feel more comfortable and suffer less of the laboratory context when performing a task at home. Third, the costs and time spent to recruit participants and collect data can be particularly lowered using internet platforms that provide automatically registered responses and calculate scores.

However, it is worth saying that collecting data using online testing needs specific exclusion criteria to maximize control (e.g., Gosling et al., 2004). In addition, online testing is subject to external interference, such as potential internet connection interruptions, and uncontrolled external sounds. Moreover, the recruiting process needs to be particularly “catchy” to convince participants to take part to the study (e.g., promising a final feedback). Finally, this method may exclude particularly old people, who are rarely able to perform an experiment alone using a device without help. However, it represents an important way to reach people who cannot move autonomously, or prefer to stay at home.

Regarding the measurement of musical abilities online, some research laboratories currently offer the possibility to perform tests in this format (e.g., Harvard and Innsbruck universities), simply using an internet connection (e.g., Profile of Music Perception Skills; Law & Zentner, 2012). These tests, proposed with an attractive layout, and with the promise of a final feedback, provide the person who completes them with some information about their performance with respect to the other participants. This format makes self-administration of the tests at home possible at any time, allowing researchers to reach people for whom it would be difficult or impossible to travel to a testing center or to the office of a testing professional.

However, it is worth mentioning that always keeping open the possibility of performing a test allows a person to perform it several times or quit without reacing to the end of the test. Though, when collecting data using an online method, it is advisable to use an online platform created for this very purpose, such as LabVanced (Finger et al., 2017) or Gorilla (Anwyl-Irvine et al., 2020), to maximize the reliability of the results. As reported by Tsuji et al. (2022), platforms such as these have demonstrated to be very effective means to collect data for experimental aims. For instance, Marimon et al. (2021) used LabVanced to collect reaction times of 3- to 8-year-old children to assess their sensitivity to nonadjacent dependencies in linguistic stimuli. In addition, Ross-Sheehy et al. (2021)

used the Gorilla platform to record button presses from 4- to 10-year-old children in a visual WM task.

Therefore, investigating music abilities using means that are more engaging and stimulating, as well as providing the possibility to perform the tests at home with a simple internet connection, is now an important option to consider in the psychology of music research field.

2. Study 1: Listening to music, affect and cognition in healthy aging

2.1 Rationale and aims of the study

Studies on aging have recently focused on developing interventions to delay age-related cognitive decline, together with promoting healthy aging. Because age-related declines in some cognitive abilities pose a threat to the health and well-being of the older people, interventions with even a modicum of success are important to document.

In the present investigation, we tested a sample of healthy older individuals and asked whether listening to brief excerpts of music had positive effects on their cognitive abilities and affective states. As discussed in Chapter 1, the effect of music listening on cognition and emotion have been studied when the music is played either (a) while participants complete a task (i.e., background music) or (b) before they begin the task, with mixed results, presumably because of large individual differences (e.g., personality) and contextual factors (e.g., task difficulty; Hallam et al., 2009; Kiss & Linnell, 2022). However, evidence shows that among healthy or pathological (e.g., Alzheimer's patients) older adults, quiet classical music played in the background enhances semantic fluency, perhaps by increasing arousal levels and attention (Thompson et al., 2005). In one instance (Mammarella et al., 2007), quiet Baroque music enhanced phonemic fluency as well as short-term memory (forward digit span) among healthy older adults. In another (Ferreri et al., 2014), lively jazz music appeared to facilitate older adults' memory for lists of words by decreasing activity in the prefrontal cortex. Older adults' memory and processing speed can also be improved when classical music is played in the background, although the improvement in processing speed is evident only when the music has a fast tempo (Bottiroli et al., 2014). Other results show no effect of music listening on younger or older adults' WM or short-term memory (Borella et al., 2014; Giannouli et al., 2019). In a test of visuospatial abilities (block design), older but not younger adults seemed to perform better after listening to music, but the advantage extended to frequency-modulated noise (Padulo et al., 2020).

In general, older adults appear to have executive-function deficits (Braver & West, 2005). Hasher and Zacks (1988) sustained that older adults fail to inhibit attending to irrelevant information during encoding and retrieval. For example, when instructed to ignore italicized text, older adults' reading speed is negatively affected by the presence of text in italics, and more so compared to younger adults (Connelly et al., 1991). From this view, background music should have a stronger negative effect on older compared to younger adults, because it would be more difficult to ignore for older people. In line with this view, older but not younger adults' associative memory is impaired in the presence of background music, even though both age groups self-report that the music is distracting (Reaves et al., 2016). Indeed, positive results among the elderly may apply only to low-amplitude background music. Among young adults, loud background music is more disruptive to reading comprehension than the same music presented at a lower volume is (Thompson et al., 2012). For this reason, we did not present music concurrently with the cognitive tasks in the study. Rather, participants completed the tasks after they listened to music. They also completed parallel versions of the same tasks before the music to obtain a baseline measure of performance as a way to monitor changes in cognitive performance as a function of music listening.

Although studies of the Mozart effect (see Rauscher et al., 1993) originally focused solely on cognitive outcomes while ignoring the mediating factor of emotional state, assuming that listening to classical music directly activates brain areas that enhance cognition, subsequent research revealed that the effect is indeed mediated by emotional state, specifically arousal and mood (Thompson et al., 2001). The arousal and mood hypothesis, discussed in the first Chapter, proposes that music can cause changes in arousal and/or mood, which in turn can influence cognitive performance. In other words, the link between music and cognition is mediated by emotion.

Considering the literature on the effect of music listening on cognitive performance, the duration of the exposure to the listening condition was hardly ever contemplated as a relevant aspect. Indeed, in the studies investigating music listening and cognitive tasks, the length of the listening condition varies from a few seconds (e.g., Silva et al., 2020), to 10 mins (e.g., Hirokawa, 2004). In addition, although some literature has found that the effect of music listening lasts for about the 12 mins after the listening condition (see Jenkins, 2001) and then gradually disappears, the duration of the music listening might have a different effect on each listener (e.g., Linneman et al., 2018), and particularly in an older listener, because older adults generally need more time to focus on and process auditory stimuli (Craik & Salthouse 2008).

Therefore, in our study, we presented two versions (a shorter and a longer one) of each listening condition, to explore whether listening to a piece for less amount of time would be the same as listening for longer, in our sample of older adults.

To our knowledge, this is the first study to test the arousal and mood hypothesis in a sample of older adults. The musical stimuli were the same as those used previously (Borella et al., 2014, 2019; He et al., 2017; Thompson et al., 2001) simply because they are clearly happy- or sad-sounding, they and have been used previously and successfully to affect arousal and/or mood, and subsequently cognition. In addition to measuring arousal and mood before and after music listening, we also measured dominance, which is sometimes considered to be a third dimension of affective responding (Bradley & Lang, 1994; Mehrabian & Russell, 1974). We hypothesized that the music manipulations would affect arousal and mood but not dominance, thereby providing a test of discriminant validity. Although Mammarella et al. (2007) tested the arousal and mood hypothesis in older adults, they did not measure emotional response and they presented the music concurrently with the cognitive tasks, and therefore attentional limitations could have affected cognitive performance. As for the control condition, because sitting completely unoccupied in silence represents an infrequent scenario in real life (Chanda & Levitin, 2013), we chose a spoken-word recording, to compare music listening with a similar reward as music in terms of level of arousal, attentional capture, and engagement.

Our dependent measures were tests of executive functions (verbal fluency, flexibility and speed), WM, and a test of arithmetic ability. We focused on executive functions because of the declines observed in older individuals, which extend to healthy aging, mild cognitive impairment, and Alzheimer’s disease (Borella et al., 2006; Braver & West, 2008; Buckner, 2004; Ferguson et al., 2021). WM, in particular, declines steadily from around age 40 into old age (Ferguson et al., 2021). Thus, manipulations that have positive effects in healthy aging might also be informative in the development of interventions for age-related cognitive decline. Because previous studies have tested fluency and short-term memory in older adults either reported null results (Giannouli et al., 2019) or presented music while participants completed the cognitive tasks (Mammarella et al., 2007), our study was largely exploratory. We also included a test of arithmetic abilities to determine whether any observed effects might extend to abilities that are required in everyday life. Because deficits in simple subtraction abilities are evident in healthy and pathological aging (Arnaud et al., 2008), it is particularly important to identify manipulations that have positive effects.

In line with the arousal and mood hypothesis, we expected to find a positive effect of listening to music on mood and arousal, and on cognitive performance. In particular, we expected to find an improvement of both mood and arousal after listening to the happy-sounding music piece, and the opposite arousal and mood trends for the sad-sounding music piece. In addition, we expected better cognitive performance after listening to the music conditions with respect to the control condition.

2.2 Method

2.2.1 Participants

Participants were 132[1] older-adult (65–75 years) volunteers. All were healthy, native Italian-speaking, community-dwelling individuals, recruited through word of mouth or from associations for the elderly in northeast and southern Italy. None was a professional musician, and all reported listening to classical music only occasionally. Inclusion criteria were good physical and mental health, assessed with a semistructured interview (De Beni et al., 2008), and good cognitive functioning, using a cutoff score of 8 on the Italian Checklist for Multidimensional Assessment (SVAMA, Gallina et al., 2006). No participant scored significantly below age- and education-

---

[1] Applying all coefficients in the models, a power analysis showed an estimated sample of 18–20 in each group should be sufficient to obtain a power of .80, an effect size of .30, and a *p* value less than .05 (using the R software's pwr library).

matched norms on the Vocabulary subtest from the Wechsler Adult Intelligence Scale – Revised (WAIS-R) (Wechsler, 1981). Prescreening also included a well-being measure (Psychological Wellbeing Questionnaire, PWB-Q, Personal Satisfaction subscale; De Beni et al., 2008), and none of the participants had scored behind the cut-off representing a low personal satisfaction. Then, and a measure of trait positive and negative affect (Positive and Negative Affect Scale, PANAS; Watson et al., 1988) was used. The Adaptive Functions of Music Listening Scale (AFML; Groarke & Hogan, 2018), which asks participants about why they listen to music, was also administered to confirm that music-listening habits did not differ across groups. Its total score represents the degree to which music listening is used adaptively (e.g., stress regulation or strong emotional experiences).

2.2.2 Materials

Auditory Stimuli

Two pieces of music, the same as those used in previous research (Borella et al., 2014, 2019; He et al., 2017; Thompson et al, 2001), were excerpted from recordings of Classical and Baroque music. None of the participants was familiar with the two pieces. One piece was characterized by a happy-sounding music: specifically, it was the first movement (*Allegro con spirito*) from Mozart's sonata K 448 for two pianos in D major[2]. As noted, the piece is in major mode with a relatively fast tempo. The other excerpt was the sad-sounding *Adagio* in G minor for organ and strings[3]. Although the piece is commonly attributed to the Baroque composer Tomaso Albinoni, it is now thought to have been composed in the mid-20th century by the Italian musicologist Remo Giazotto (Talbot, 1971). It is in minor mode and has a slow tempo.

---

[2] Mozart, W.A. (1985). Sonata for two pianos in D major, K 448 (K. 3375a) [Recorded by M. Perahia & R. Lupu]. On Music for piano, four hands [CD]. London: Sony Classical. (1992). See Appendix 1 for the complete score.

[3] Albinoni, T.G. (1981). Adagio in G minor for organ and strings [Recorded by I Solisti Veneti, conducted by C. Scimone]. On Albinoni's adagios [CD]. Perivale, England: Warner Classics. (1996). See Appendix 2 for the complete score.

We used Audacity software to create two versions of both excerpts: a longer one (approwimately 8.2 min) and a shorter one (approximately 2.5 min). The excerpts were transferred digitally from CD without loss of sound quality (44.1 kHz, 16 bit). For the Mozart sonata, which is composed in typical Sonata form, the longer version comprised the first movement of Mozart sonata K 448, approximately the same duration (8:31) as the longer version of the Albinoni *Adagio* (8:16). For the shorter versions, we used the first refrain of the Mozart sonata K 448, and two repetitions of the Albinoni *Adagio* main theme with final fadeouts (see Appendix 1 and 2 for the complete scores, highlighting where the cuts had been made).

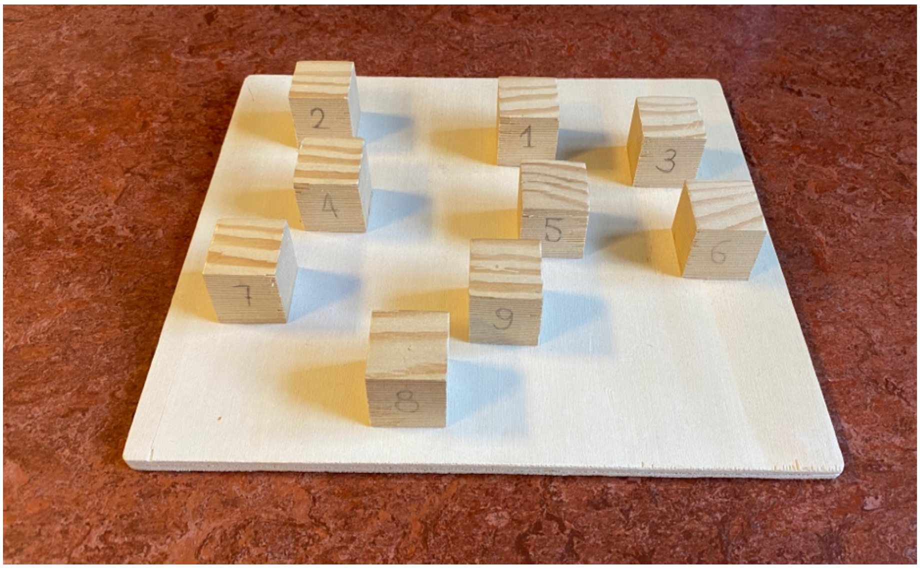
Short and long versions of the auditory stimulus in the control condition were similar in duration and meant to be approximately as engaging as listening to the music (Borella et al., 2014). They comprised a short description of the invention of the television, adapted from a standardized test of listening comprehension (Carretti et al., 2013).

Measures of Cognition and Affect

The cognitive tasks measured executive functions (verbal fluency as well as cognitive flexibility and speed) and visuospatial WM. An additional test measured arithmetical ability. Two versions of each task were created to be of equal difficulty, so that they could be counterbalanced with pre- and posttesting.

Visuospatial WM was tested with the backward Corsi blocks task (adapted from Corsi, 1972). In this task, participants are presented with nine blocks arranged randomly on a wooden tablet. The experimenter taps a sequence of blocks and asks the participant to tap the same blocks but in reverse order. Tap sequences increase in length (from two to seven), with one trial for each sequence length. The dependent variable was the longest sequence completed successfully (maximum six; see Figure 1 for a photo of the wooden table used).

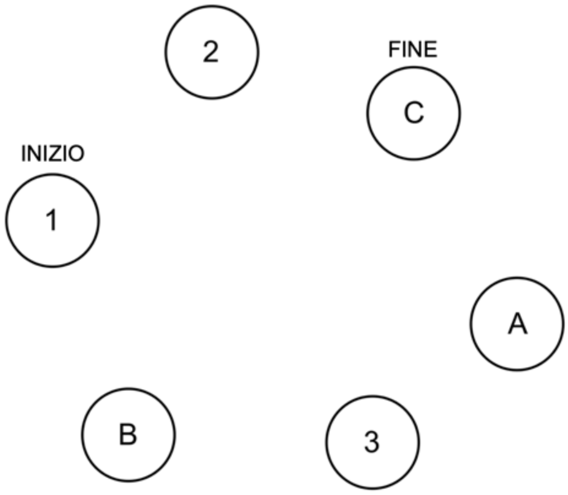
Figure 1. *Corsi blocks task. The participants sits in front of the experimenter, who is the only one seeing the numbers.*



The measure of verbal fluency was adapted from Novelli et al. (1986). Participants were given 1 min to generate as many words as possible that begin with a given letter (F or P), excluding proper names. Scores were the total number of appropriate words the participant produced orally.

Cognitive flexibility and speed were tested with the Trail Making Test B (TMT-B; Amodio et al., 2002). Participants were presented with a sheet of paper with 25 circles (diameter 2 cm) containing numbers from 1 to 13 or 12 letters from A to N, all in the same font (24-point Arial). M and N were substituted for J and K of the original test because the Italian alphabet does not include J and K. Participants used a pencil to connect the circles in sequential order, alternating between numbers and letters (1–A–2–B–3–C, and so on) as accurately and rapidly as possible. The time taken to complete the task was used as the dependent variable (see Figure 2 for the example trial).

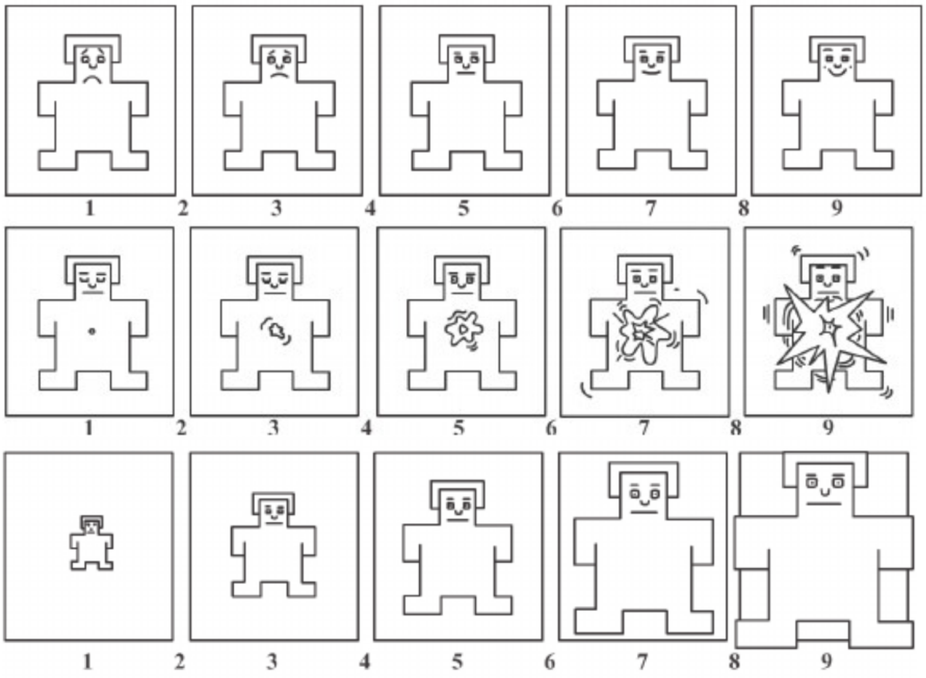
Figure 2. *TMT-B example trial.*



Arithmetic ability was measured with the AC-FL (*Prove di fluenza nelle abilità di calcolo per il secondo ciclo della scuola primaria*; Caviola, et al., 2016), which measures the speed and accuracy with which participants can mentally add, subtract, and multiply (blocked presentation). Here, we used only the addition and subtraction blocks, administered in that order. Participants had 1 min to complete as many operations as possible. The dependent variable was the total number of correct additions and subtractions.

Finally, we measured arousal and mood (valence) with the Self-Assessment Manikin (SAM; Bradley & Lang, 1994), a questionnaire that comprises pictures taken from the International Affective Picture System (Lang et al., 1988). On each trial, participants viewed a single picture and rated their felt emotional response using three 9-point scales (arousal, mood, and dominance). For each scale, 1 indicated the lowest rating (e.g., low arousal) and 9 indicated the highest (e.g., high arousal) (see Figure 3).

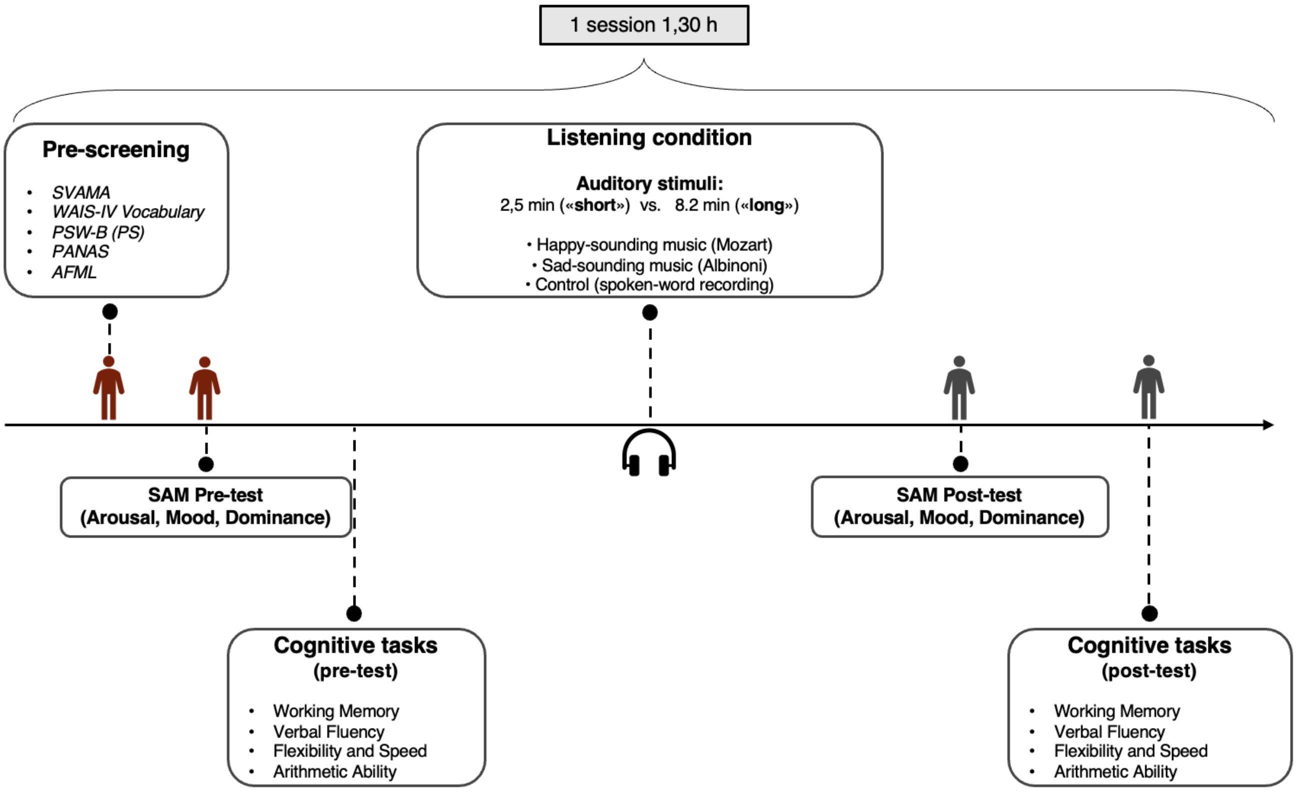
Figure 3. *SAM mood, arousal and dominance visual representation (Bradley & Lang, 1994).*



2.2.3 Procedure

Participants were tested in a single 90-min session (see Figure 4). After the prescreening tests (SVAMA, WAIS-IV Vocabulary, PSW-B, PANAS, and AFML in that order), they completed the SAM followed by the cognitive tasks in the following order: TMT-B, backward Corsi blocks, AC-FL, and verbal fluency. After completing the cognitive tasks, participants were assigned randomly to one of six listening conditions formed in a 3 × 2 factorial design based on the music (Mozart, Albinoni, and control) and duration (short and long). Participants listened over headphones (Sennheiser HD 280 Pro) while the experimenter watched and ensured that they were not distracted. After the listening session, participants completed the SAM again, followed by the parallel versions of the cognitive tasks in the same order.

Figure 4. *Visual representation of the experimental procedure adopted in Study 1.*



2.3 Results

In order to confirm that participants assigned to the six different listening conditions did not differ at pre-test (i.e., before they listened to any listening condition), considering that our data were normally distributed, we used one-way between-subjects ANOVAs to test for differences in demographics (age, education) and the pre-screening measures (SVAMA, Vocabulary, PWB-Q, PANAS-Positive, PANAS-Negative, AFML). No significant group differences emerged, *ps* ≥ .08. Descriptive and inferential statistics are provided in Table 4.

The following analyses consist of: (1) group comparisons on the cognitive tasks, (2) group comparisons on the emotion measures, and (3) associations between cognition and emotion scores, specifically those that showed reliable group differences in the previous analyses.

Table 4. *Statistical Comparisons of the Six Groups of Participants at Pre-Test (ns = 22).*

	Short Version						**Long Version**							
	Mozart		**Albinoni**		**Control**		**Mozart**		**Albinoni**		**Control**			
	M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***F*(5, 126)**	***p***
PRE-SCREENING														
Age	69.77	3.25	69.68	3.01	69.32	2.98	68.59	3.10	69.46	3.31	69.50	2.92	< 1	.842
Education	10.05	3.95	10.05	4.05	10.73	3.52	10.55	4.03	11.23	4.48	10.77	4.34	< 1	.925
Vocabulary	42.86	11.95	40.46	12.27	45.32	10.75	39.50	10.43	41.05	9.60	38.64	12.31	< 1	.438
PWB-PS	33.66	4.09	34.05	5.22	33.25	5.09	34.46	5.48	34.55	5.31	34.32	5.89	< 1	.557
PANAS-P	32.02	4.94	32.14	6.57	31.09	5.22	33.89	6.42	32.5	6.97	32.93	6.55	1.27	.282
PANAS-N	18.14	5.28	18.34	6.18	18.09	5.14	18.5	5.39	19.68	5.57	19.84	5.58	< 1	.571
AFML	146.91	24.09	150.14	19.06	150.50	22.90	152.86	12.71	153.00	11.54	154.23	15.85	< 1	.805
COGNITION														
Working Memory	2.73	1.24	2.59	1.10	2.50	0.96	2.32	0.57	2.82	1.01	2.64	1.00	< 1	.640
Flexibility/Speed	118.50	20.11	117.96	30.74	119.82	23.61	122.05	31.60	123.73	32.03	124.09	35.98	< 1	.971
Verbal Fluency	11.91	4.02	13.05	3.48	12.46	3.85	12.55	3.84	12.68	4.01	12.68	3.84	< 1	.878
Arithmetic	38.91	10.94	38.59	7.06	39.18	6.32	36.27	11.27	37.86	9.53	36.82	10.51	< 1	.887
EMOTION														
Arousal	4.59	1.41	5.05	2.28	4.73	2.05	4.27	2.51	4.18	2.34	4.09	1.72	< 1	.638
Mood	6.55	0.96	6.46	1.37	6.14	0.83	5.18	1.22	6.36	1.22	5.96	1.25	< 1	.455
Dominance	6.64	1.40	7.18	1.33	6.91	2.09	7.05	1.68	6.59	1.62	6.91	1.27	< 1	.807

Cognitive Tasks

Preliminary analyses confirmed that there were no group differences at pre-test on the four cognitive tasks, *ps* > .6. Separate mixed-design ANOVAs with one repeated measure (pre- or post-test) and two between-subjects factors (listening condition, duration) were conducted for each of the four cognitive variables. Means and SDs are provided in Table 5.

Table 5. *Descriptive Statistics (M and SD) for Measures of Cognition and Emotion at Pre- and Post-Test.*

		Short version						**Long version**					
		Mozart		Albinoni		Control		Mozart		Albinoni		Control	
		M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
COGNITION													
Working Memory	Pre	2.73	1.24	2.59	1.10	2.50	0.96	2.32	0.57	2.82	1.01	2.64	1.00
	Post	3.32	1.52	2.68	1.49	2.50	0.67	2.91	0.91	2.86	0.89	2.68	0.78
Flexibility/Speed	Pre	118.50	20.11	117.96	30.74	119.82	23.61	122.05	31.60	123.73	32.03	124.09	35.98
	Post	89.59	27.50	95.23	34.95	106.27	19.88	102.18	29.96	104.00	33.23	102.64	31.16
Verbal Fluency	Pre	11.91	4.02	13.05	3.48	12.46	3.85	12.55	3.84	12.68	4.01	12.68	3.84
	Post	12.86	3.63	13.00	2.33	14.05	4.12	13.05	5.12	13.59	3.51	13.86	3.12
Arithmetic	Pre	38.91	10.94	38.59	7.06	39.18	6.32	36.27	11.27	37.86	9.53	36.82	10.51
	Post	41.36	9.95	40.68	5.31	41.64	5.31	39.05	10.25	40.36	7.31	38.64	9.16
EMOTION													
Mood	Pre	6.55	0.96	6.46	1.37	6.14	0.83	5.18	1.22	6.36	1.22	5.96	1.25
	Post	6.68	0.89	6.22	1.34	5.86	0.99	7.82	0.96	6.59	1.30	5.73	1.12
Arousal	Pre	4.59	1.41	5.05	2.28	4.73	2.05	4.27	2.51	4.18	2.34	4.09	1.72
	Post	5.14	1.91	3.09	1.41	4.59	2.04	5.36	2.48	3.32	2.17	4.09	1.44
Dominance	Pre	6.64	1.40	7.18	1.33	6.91	2.09	7.05	1.68	6.59	1.62	6.91	1.27
	Post	6.59	1.37	6.55	1.47	7.05	1.21	7.14	1.75	6.50	1.57	6.95	1.53

Note. All ns = 22; See text for details of the specific measures.

For WM, there was a main effect of test session, with improvement from pre- to post-test, $F_{(1, 126)} = 11.25$, $p < .001$, partial $\eta^2 = .082$. This main effect was qualified, however, by a two-way interaction between test session and music, $F_{(2, 126)} = 7.24$, $p = .001$, partial $\eta^2 = .103$. There were no other main effects or interactions, *ps* > .2. Follow-up tests of the interaction revealed that although there was a large pre-to-post improvement in WM for the Mozart group, $p < .001$, there was no change for the Albinoni, $p = .263$, or control, $p = .508$, group.

Analysis of verbal fluency scores revealed an improvement from pre- to post-test, $F_{(1, 126)} = 8.36$, $p = .005$, partial $\eta^2 = .062$, but no other main effects or interactions, *ps* > .5. The results were identical for mental flexibility and speed, with a large improvement (i.e., faster performance) from pre-to post-test $F_{(1, 126)} = 140.63$, $p < .001$, partial $\eta^2 = .527$, but no other main effects or interactions, *ps* > .1; and for arithmetic, with a substantial improvement over time, $F_{(1, 126)} = 42.56$, $p < .001$, partial $\eta^2 = .252$, but null results otherwise, *ps* > .2.

Emotion Tasks

We used the same analytic method for emotional responses collected at pre- and post-test with the SAM. Means and SDs are provided in Table 5. For Arousal scores, a marginal effect of music, $F_{(1, 126)} = 3.01$, $p = .053$, partial $\eta^2 = .046$, was qualified by a two-way interaction between music and test session, $F_{(2, 126)} = 15.37$, $p < .001$, partial $\eta^2 = .196$. Arousal levels increased after listening to Mozart, $p = .009$, but decreased after listening to Albinoni, $p < .001$. They did not change for participants in the control condition, $p > .5$. There were no other main effects or interactions, *ps* > .07.

As for Mood scores, a significant main effect of music, $F_{(1, 126)} = 11.09$, $p < .001$, partial $\eta^2 = .150$, was qualified by a two-way interaction between music and test session, $F_{(2, 126)} = 7.86$, $p < .001$, partial $\eta^2 = .111$. Follow-up tests revealed a significant improvement in mood for the Mozart group, $p < .001$, but no improvement for the Albinoni group, $p > .9$, and a marginal decrease for the control group, $p = .057$. All other tests were non-significant, *ps* > .09, except for a two-way interaction between duration and test session, $F_{(1, 126)} = 7.15$, $p = .008$, partial $\eta^2 = .054$. Improvements in positive mood were evident for participants in the long-duration conditions, $p = .002$, but not for those in the short-duration conditions, $p = .444$. There were no main effects or interactions in the analysis of Dominance scores, *ps* > .3.

Associations Between Cognition and Emotion

Finally, we asked whether improvements in WM in the Mozart condition were associated with changes in arousal and mood. Improvements in WM, arousal, and mood were calculated as difference scores (post – pre). Increases in arousal tended to be ac-companied by improvements in positive mood, $r = .298$. N = 132, $p < .001$. Nevertheless, changes in WM were not associated with increases in arousal, $p > .3$, or with improvements in mood, $p > .2$. The lack of an association between changes in WM and changes in arousal or mood precluded the possibility of mediation analysis. Moreover, when we used a general linear model to predict improvements in WM as a function of music, increases in arousal, and improvements in mood, the effect of music remained significant, $F_{(2, 127)} = 6.35$, $p = .002$, partial $\eta^2 = .091$, with larger improvements in WM for the Mozart group than for the Albinoni group, $p = .011$, and the control group, $p = .004$, but no difference between the Albinoni and control group, $p > .9$ (Tukey's test). Increases in arousal, $p > .8$, and improvements in positive mood, $p > .9$, had no partial association with improvements in WM when group differences were held constant.

The lack of associations between improvements in WM and changes in arousal or mood is based on a failure to reject the null hypothesis, meaning that the observed data are not particularly unlikely if the true association is actually null. To explore further this null result, we used Bayesian analyses conducted with JASP 0.16.1 (JASP Team, 2022; default priors) to test whether the observed data were more likely under the null (no association) or alternative (association) hypothesis. For the association between changes in WM and increases in arousal, the observed data were 5.59 more likely under the null than the alternative hypothesis. For the association between changes in WM and improvements in mood, the observed data were 4.81 times more likely under the null hypothesis. In other words, the data provided substantial evidence for the null hypothesis in both instances (Kass & Rafferty, 1995). By contrast, for the association between increases in arousal and improvements in mood, the observed data were 42.5 times more likely under the alternative hypothesis.

2.4 Discussion

In this study, we examined whether happy-sounding music improved the emotional state of active, autonomous, typically aging older adults, and whether such improvement was accompanied by improvement in cognitive performance. As expected, listening to happy-sounding music (Mozart's sonata K 448) increased arousal levels and improved mood, compared to listening to sad-sounding music (Albinoni's *Adagio*) or to control conditions that involved a spoken-word recording. However, listening to sad-sounding music was accompanied by reduced levels of arousal. As predicted, no effect of music listening on dominance levels occurred, but participants who heard happy-sounding music exhibited improvements in visuospatial WM. Therefore, at group level, the results are consistent with the arousal and mood hypothesis. At the individual level, however, improvements in arousal and/or mood had no association with improvements in WM. In fact, the observed data provided substantial evidence for no association.

These apparently contradictory results could be explained considering that for younger adults, cognitive benefits emerge reliably after listening to happy-sounding music, and these benefits are a consequence of changes in arousal and/or mood (He et al., 2017; Husain et al., 2002; Thompson et al., 2001). The greater individual differences among older adults make, in some cases, correlations between scores on one task and scores on another relatively difficult to observe. In fact, previous mood-induction research has highlighted the importance of considering age-related factors (Larcom & Isaacowitz, 2009). Older adults who regulate their emotions slowly are more likely than their faster-regulating counterparts to exhibit elevated levels of anxiety and symptoms of depression, but less likely to be optimistic (Laukka & Juslin, 2017). In addition, they retain a positively induced mood for a shorter period, comparable to younger adults, but not like faster-regulating older adults.

Moreover, emotional responses to music show age-related changes, such that older adults have stronger emotional reactions to happy-sounding music compared to music that conveys other emotions (Vieillard et al., 2012). They also have difficulty in recognizing negative emotions, in music as well as in speech, although the ability to recognize positive emotions in both domains remains relatively stable throughout adulthood (Lima & Castro, 2011; Laukka & Juslin, 2007). In some cases, older adults show recognition deficits for positive and for negative emotions, for music as well as for faces (Sutcliffe et al., 2017). However, such deficits appear to be larger for music that conveys negative emotions (Pearce & Halpern, 2015). In this, the "positivity effect" refers to findings showing that older individuals in general preferably attend to and remember positive compared to negative stimuli (Carstensen & DeLiema, 2018). These factors could explain why older adults in our study did not show a decrease in mood after listening to the sad-sounding music, even though they had lower arousal levels.

Reviews and studies of reaction-time data reveal greater intraindividual differences in typical (Fagot et al., 2018) as well as pathological aging (Haynes et al., 2017). Trial-by-trial fluctuations in individual participants' responses for make response times inherently noisy when measured among older adults. The lack of an association between emotional response and cognitive ability observed here could also stem from differences in measurement approaches, with our WM task being more effortful than self-reports of emotional responding (Olderbak et al., 2017). Nevertheless, an association between other self-report emotion tasks and objective cognitive measures has been observed previously with younger adults (He et al., 2017; Husain et al., 2002; Thompson et al., 2001). In other words, the effort required to do a task appears to be moderated by age.

A final possibility is that mood induction has a weaker or no association with cognitive performance in older compared to younger adulthood. For instance, film clips were used to induce positive or negative moods, and performance on a subsequent test of executive function (i.e., Tower of London, requiring a specific number of moves to be done) was impaired for older adults in both mood-induction conditions, compared to a neutral control condition (Philips et al., 2002), whereas young adults' performance was similar across conditions. This finding could explain why performance in the WM task improved, whereas it did not on the two executive function tasks. Some planning is obviously required for good performance on the Trail Making Test-B, as in the Tower of London task, whereas it is less involved in the verbal-fluency task. In addition, many different executive function tasks exist, and it seems unlikely that mood induction for older adults would have similar detrimental effects for all of them. However, because we did not compare our sample with one of young adults, further research is needed to substantiate this hypothesis. Among older adults, WM is known correlate highly with many aspects of executive functioning (McCabe et al., 2010). In fact, differential performance on the WM task—but not on the other cognitive tasks—might be related to the fact that only the WM test was not timed. Scores on the test of cognitive flexibility and speed (TMT-B) were calculated as the time taken to complete the task, and also scores on the verbal-fluency and arithmetic tests were calculated as the number of correct responses in 1 min, relying on speed in responding. In general, responses begin to slow down after 20 years of age; they become even slower over the age of 60, due to actual slowing of mental processing speed (von Krause et al., 2022). Additionally, the WM task may also have been more difficult in terms of cognitive requirement and attentional control compared to the other tasks, therefore allowing for more improvement. In any event, task-specificity remains an unresolved issue, for performance after mood induction in general, and for the so-called Mozart effect in particular (Schellenberg et al., 2012).

We also investigated whether effects of music listening would change varying the duration of music stimuli (i.e., short or long). Compelling evidence would have come from a significant three-way interaction, indicating that changes from pre- to posttest were evident only for listeners who heard the long-duration Mozart excerpt. This interaction was not significant. Nevertheless, Bayesian analyses documented positive evidence of improvements in WM for listeners in the Mozart group: strong evidence for the short-duration group, and even stronger evidence for those who heard the longer duration, emerged. In addition, improvements in mood were evident only for participants in the Mozart-long condition. Although these findings need to be replicated in future research, our results do not allow us to recommend using the shorter-duration excerpts for applied or research purposes, instead, the longer duration seems to be better for older adults. Indeed, because older people generally need more time to process auditory stimuli (Craik & Salthouse, 2008), and perhaps to feel the effect, they might benefit from longer exposure to the music. This speculation needs to be tested in future studies.

To conclude, older adults who heard the happy-sounding music exhibited larger positive changes in arousal, mood, and WM, a finding predicted by the arousal and mood hypothesis (Thompson et al., 2001). Nevertheless, the independence between measures of emotion and cognition appears to stem from large inter- and intraindividual variability that characterizes the entire aging process.

3. Study 2: Music training, personality and cognitive abilities: professional musicians across the adult life span[4]

3.1 Rationale and aims of the study

Musical training has positive associations with a variety of cognitive abilities (see Swaminathan & Schellenberg, 2019 for a review). As previously discussed in Chapter 1, however, most of this evidence comes from correlational and quasi-experimental studies, precluding inferences of causation. (Schellenberg, 2020).

Also, musical training is associated with demographic, personality, and cognitive variables during childhood (Corrigall et al., 2013). These preexisting individual differences in musical and nonmusical variables make musically trained individuals a poor model for the study of transfer or plasticity, despite some researchers claims to the contrary (e.g., Steele & Zatorre, 2018).

In the present investigation, our primary focus was on individuals with the highest levels of musical experience: professional musicians. Here we defined (1) professional musicians as those whose careers involve music instruction (e.g., music professors) or performance (e.g., members of orchestras), or full-time study at the tertiary level or higher, and (2) musically trained individuals as those who had at least 6 years of lessons and were not working as musicians. As discussed in Chapter 1, the "six-year rule", although generally accepted as a threshold for musical expertise (Zhang et al., 2020), typically ignores whether individuals are working as musicians. This issue is particularly important because of findings showing that music training, when treated as a continuous variable (i.e., duration of formal lessons), has a positive linear association with general cognitive ability in childhood and in adulthood (e.g., Corrigall et al., 2013; Degé et al., 2011; Swaminathan et al., 2017).

---

[4] This study has already been published: Vincenzi, M., Correia, A. I., Vanzella, P., Pinheiro, A. P., Lima, C. F., & Schellenberg, E. G. (2022). Associations between music training and cognitive abilities: The special case of professional musicians. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000481

The primary goal of the present investigation was to test our hypotheses that professional musicians are different from musically trained individuals in musical ability and personality traits, but not in cognitive abilities.

Evidence show that personality predicts occupational choices (Holland, 1997). Openness-to-experience Big Five trait, in particular, seems to be positively associated with creativity (Feist, 1998, 2019). Extraversion is another personality trait that predicts creativity (Feist, 2019). Because the Big Five traits are intercorrelated, metatraits (higher-order personality factors) have been proposed (DeYoung, 2006). Shared variance between openness and extraversion forms one metatrait that indexes behavioral engagement; shared variance among agreeableness, conscientiousness, and neuroticism forms a second metatrait indexing stability. Because engagement is an aggregate of extraversion and openness, it has a strong positive association with creativity (Feist, 2019), which extends to objectively measured creative achievements and everyday creative behaviors, including music (Sylvia et al., 2009). In a previous study, Kuckelkorn et al. (2021) compared the personalities of professional musicians to those of amateur musicians and nonmusicians. Professional musicians had higher levels of openness than amateurs, who had higher levels than nonmusicians, as one might expect, although neuroticism unexpectedly showed the same pattern. The other main finding was that, in both musician groups, singers were more extraverted than instrumentalists, except for percussionists. One problematic aspect of this study was that amateur musicians were classified as individuals who had played a musical instrument (including voice) at any point in their lives for any amount of time but were not professionally active. In other words, professional and amateur musicians differed markedly in music training as well as professional status, which makes these response patterns difficult to interpret.

In the present study, we examined group differences in musical ability, personality, and general cognitive ability in a sample that comprised professional musicians and participants who were musically trained or untrained. We expected to find robust group differences in measures of musical ability (professionals, trained, untrained). For personality, previous findings allowed us to be relatively confident that the professional and trained groups would score higher than the untrained group on openness and extraversion, and on engagement more generally. We also expected that professional musicians would have particularly high scores on these personality variables. Finally, although musically trained participants should perform better than untrained participants on a measure of general cognitive ability, we did not expect the professionals to outperform the trained group.

3.2 Method

3.2.1 Participants

The sample comprised 642 volunteer participants, who ranged in age from 18 to 84 years (M = 34.8, SD = 15.1; 384 women, 258 men), recruited primarily through social-media postings for an online study on personality and musical abilities, which was open to individuals with any level of musical expertise. To increase the study's appeal, the posting specified that participants would receive feedback about their musical ability and personality. The study was made available in four languages (Italian: n = 290, European Portuguese: n = 151, Brazilian Portuguese: n = 150, and English: n = 51), which reflected the make-up of the research team while maximizing sample size and diversity.

The sample was restricted to respondents who fell into one of three groups (see Table 6 for descriptive statistics of demographic variable for each group):

- ***professional musicians***: (n = 176) had a music-related job and/or were enrolled as students in a university-level music program;
- ***musically trained participants*** who were not professionals: (n = 121) had at least 6 years of music lessons but did not meet the criteria for professionals; this group included many amateur musicians;
- ***musically untrained participants***: (n = 345) had a maximum of 2 years of music training.

An additional 118 participants with 3–5 years of music lessons were tested but excluded because they could not be identified clearly as trained or untrained. Five other participants were tested but excluded from analyses because of self-reported poor hearing ability (n = 1) or unspecified gender (n = 4). Professional musicians were employed as music professors (n = 126), orchestral musicians (n = 41), soloists (n = 50), conductors (n = 12), choristers (n = 8), pianists (n = 26), composers (n = 25), and members of small musical ensembles (n = 67), but these categories were not mutually exclusive. A post-hoc sensitivity analysis conducted with G* Power 3.1 (Faul et al., 2007) confirmed that a sample of 642 participants had 80% power to detect small associations of $.01 < \eta^2 < .02$ (Analysis of Covariance, three covariates, $\alpha = .05$).

3.2.2 Materials

All tasks and questionnaires were adapted for Gorilla experiment Builder (Anwyl-Irvine et al., 2020), a flexible platform for online behavioral research. Each measure in the testing protocol was created originally in English. Whenever available, published translations were used for the European Portuguese, Brazilian Portuguese, and Italian versions of the tests. Otherwise, *ad hoc* translations were created by native speakers who were also fluent in English.

Questionnaires

Goldsmiths Musical Sophistication Index (Gold-MSI, Lima et al., 2020; Müllensiefen, et al., 2014). This questionnaire has 38 items that evaluate different behaviors related to music (e.g., I spend a lot of my free time doing music- related activities). The items are mixed in terms of order of presentation. For scoring purposes they are grouped to form five subtests: Active Engagement (9 items), Perceptual Abilities (9 items), Music Training (7 items), Singing Abilities (7 items), and Emotions (6 items). A General Musical Sophistication factor is also calculated from 18 items that are representative of the five subtests. For the first 31 items, participants judge how much they agree with each statement on a seven-point rating scale (1 = completely disagree, 7 = completely agree). For the final seven items, participants select one of seven alternatives from an ordinal scale that varies from item to item. For example, the scale for the statement *I listen attentively to music for ...* had options ranging from 1 (0 - 15 min per day) to 7 (4 hours or more per day). For European-Portuguese participants, was created an online version of a published translation of the Gold- MSI that has good psychometric properties (Lima et al., 2020). For the Italian translation, items from the original English version were translated to Italian independently by two translators, both of whom were native speakers of Italian, fluent in English, experienced in translating questionnaires, and experts in the psychology of music. The goal was conceptual equivalence rather than a literal translation. Discrepancies between translations were solved by discussion to create a single version, which was, in turn, evaluated by two independent colleagues for clarity of expression and whether the translation from English was appropriate. The Italian version was then back-translated by a native speaker of English who was fluent in Italian and a scholar of psychology and music. Inconsistencies between the back-translation and the original Gold-MSI were discussed and resolved among the three translators, who also consulted with two additional experts from the discipline. Finally, 10 participants completed the Italian translation of the Gold-MSI and confirmed that the items were clear. For the Brazilian-Portuguese version, a native speaker, who was also fluent in English and an expert in the psychology of music, made minor modifications to the European- Portuguese version. Cronbach's alphas for the entire sample and for the previously unpublished (Italian and Brazilian-Portuguese) translations of the Gold-MSI were calculated. Internal reliability was maintained for the previously unpublished translations.

Big-Five Inventory (BFI) (John & Srivastava, 1999). It is a self-report questionnaire including 44 items, which measure the traits from the five-factor model of personality (McCrae & John, 1992): Openness-to-Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Participants rate how well each item describes them on a 5-point rating scale. The five personality traits are calculated as mean scores. Metatraits scores are derived by using principal-components analysis to extract the shared variance between openness and extraversion scores to form *engagement* scores, and the shared variance among agreeableness, conscientiousness, and neuroticism scores to form *stability* scores.

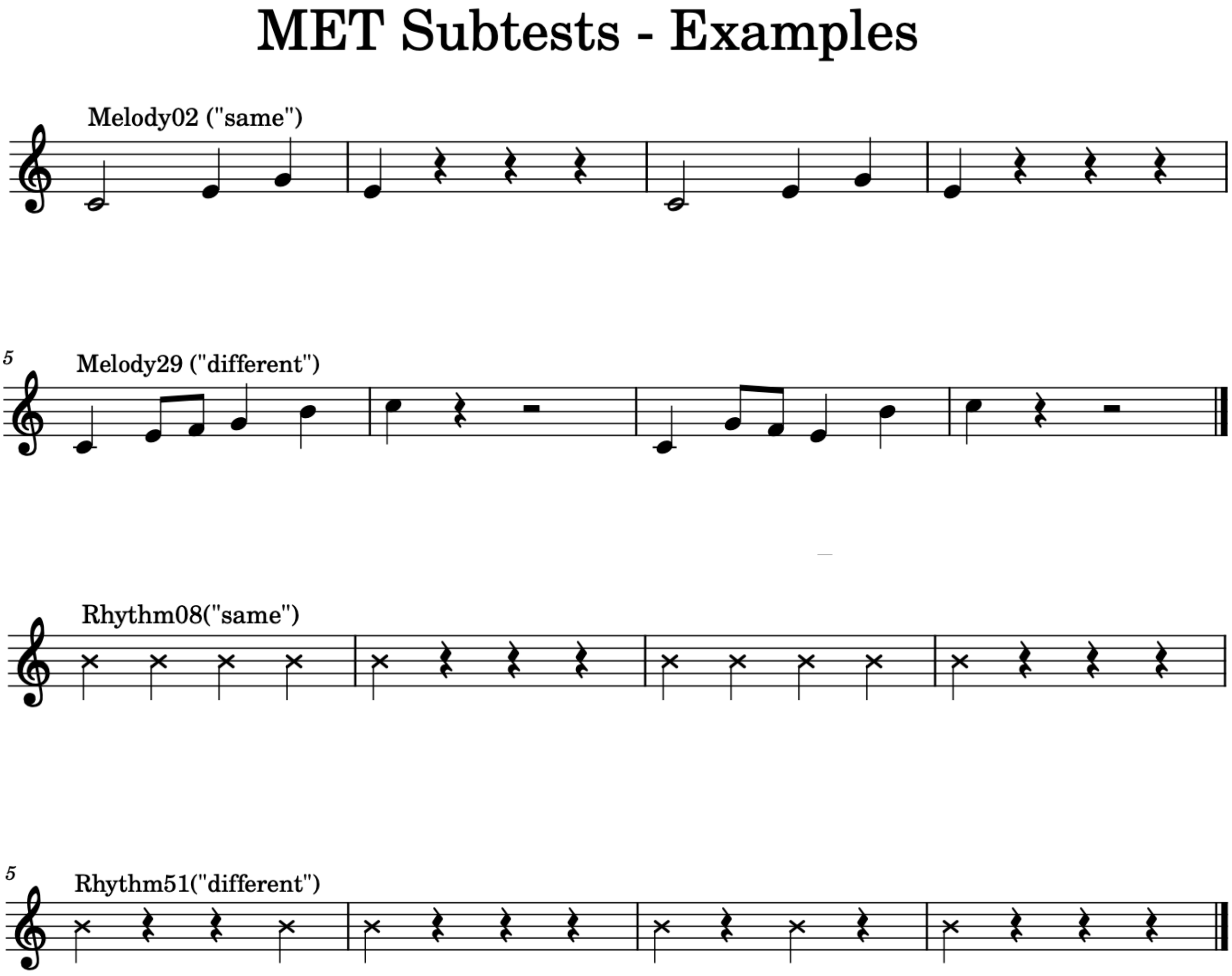
Mind-Wandering Questionnaire (MWQ, Mrazek et al., 2013). It is a 5-item questionnaire measuring trait levels of mind-wandering (e.g., *I find myself listening with one ear, thinking about something else at the same time*). Participants rate their agreement with each item on a 6-point rating scale (1 = almost never, 6 = almost always). The mean serves as an index of an individual's frequency of mind-wandering.

Objective Musical Ability

Musical Ear Test (MET, Wallentin et al., 2010). The MET is an objective measure of musical ability that has two subtests, Melody and Rhythm, presented in that order. On each of 52 trials per subtest, participants hear two short sequences of piano tones (Melody) or drumbeats (Rhythm) and judge whether they are or not identical. Half of the trials are different, such that one or more tones are displaced in the Melody subtest, and one or more beats are altered in the Rhythm subtest. Scores for both subtests are calculated as the number of correct responses.

Since the MET was administered at the end of the testing session and was relatively long (approximately 20 min), some participants did not finish it, or provided incomplete data. So, MET (Melody or Rhythm scores) were excluded for participants with more than 10 (of 52) or 5 consecutive missing responses on a subtest. Sample sizes were therefore smaller when analyses included the Melody (n = 546) or the Rhythm (n = 529) subtest. See Figure 5 for MET examples in musical notation.

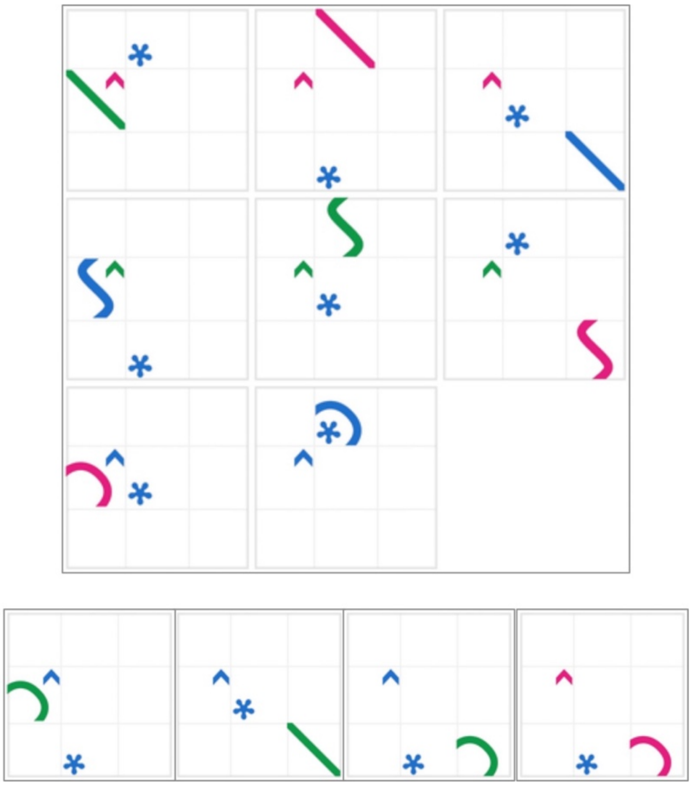
Figure 5. *Examples of the Melody and Rhythm subtests of the MET.*



General Cognitive Ability

Matrix Reasoning Item Bank (MaRs-IB, Chierchia et al., 2019). The MaRs-IB is a freely available online measure of abstract (nonverbal) reasoning, modeled after Raven's Advanced Progressive Matrices (Raven, 1965). The test has 80 trials. On each trial, a 3 x 3 matrix is presented. Eight of nine cells contain abstract shapes (varying on four dimensions: color, size, shape, and location), but the cell in the bottom-right corner is always empty. Participants' task is to choose one of four alternatives to complete the matrix, following the rules that govern differences among the other eight cells. The duration of the task is fixed at 8 min, but participants are not informed of the duration or the number of trials, only that they have up to 30 s to complete each trial. If participants complete the 80 trials in less than 8 min, the trials begin again in the same order, but responses from the second round are not considered in calculating scores, which are the proportion of responses given by the participant that are correct (excluding responses made in > 250 ms). Proportions were logit-transformed for statistical analyses. See Figure 6 for a visual example al the MaRs-IB.

Figure 6. *MaRs-IB example (taken and available from Gorilla platform).*



3.2.3 Procedure

After providing informed consent, participants completed the questionnaires in the following order: MWQ, Gold-MSI, and BFI. After the questionnaires, they completed the MaRs-IB. Then, the MET. At the end of the session, participants were provided with summary feedback about their personality, musical sophistication, and musical abilities. Ethical considerations precluded feedback about cognitive ability.

3.3 Results

We initially compared our three groups of participants in terms of basic demographic variables. Descriptive and inferential statistics are provided in Table 6.

Table 6. *Descriptive and Inferential Statistics for Demographic Variables.*

	Musically trained		**Musically untrained**		**Professional musicians**		**Group comparison**	
	M	*SD*	*M*	*SD*	*M*	*SD*	$F_{(2,639)}$	η^2
Age	32.40	14.85	32.70	14.18	40.94	14.48	21.39	.063
Education	3.98	1.05	4.21	1.04	4.52	0.92	16.24	.048
Gender	M/F	%M	M/F	%M	M/F	%M	$\chi2(2)$	φ
	123/222	35.6	41/80	33.9	94/82	53.4	17.75	.166

Note. All ps < .001.

Analysis of Variance (ANOVA) uncovered group differences in both age and education. Follow-up pairwise comparisons (Tukey's HSD) revealed that professional musicians were older than trained and untrained participants, *ps* < .001, who did not differ, *p* = .979. Professional musicians also had more education than trained participants, *p* = .032, and untrained ones, *p* < .001, who did not differ, *p* = .079. A chi-square test of independence indicated that the gender ratio also differed across groups, with a greater proportion of males among professional musicians than among trained participants, *p* < .001, and untrained ones, *p* < .001, who did not differ, *p* = .726. Thus, age, education, and gender were included as covariates in the statistical analyses that follow. As one would expect from the available literature (e.g., Hartshorne & Germine, 2015; Salthouse, 2009), cognitive ability also had a small negative correlation with age, *r* = -.089, N = 642, p = .023, a positive correlation with education, *r* = .190, N = 642, *p* < .001, but no association with gender, *p* = .165.

Analysis of Covariance (ANCOVA) confirmed that our three groups of participants differed on each of the music variables. Descriptive and inferential statistics are provided in Table 7.

Table 7. *Descriptive and Inferential Statistics for the Musical Ear Test (MET) and the Goldsmiths Musical Sophistication Index (Gold-MSI)*

	Musically trained		**Musically untrained**		**Professional musicians**		**Group comparison**		
	M	*SD*	*M*	*SD*	*M*	*SD*	$F_{(2,639)}$	*d.f.*	η^2
MET									
Melody	34.74	6.43	39.99	5.70	43.10	4.71	87.14	540	.241
Rhythm	36.71	5.71	39.46	5.08	40.30	4.46	18.65	523	.065
Gold-MSI									
Active engagement	3.88	1.23	4.58	0.99	5.25	0.82	97.02	636	.230
Perceptual abilities	4.94	1.07	6.03	0.69	6.38	0.55	165.58	636	.341
Music training	2.39	1.19	5.36	0.78	6.05	0.55	923.47	636	.341
Singing abilities	3.76	1.35	4.97	0.96	5.41	0.89	132.18	636	.292
Emotions	5.49	0.95	6.00	0.67	6.05	0.77	41.66	636	.112
General factor	3.55	1.12	5.06	0.77	5.73	0.52	352.50	636	.521

Note. All ps < .001. Age, education, and gender were held constant in the group comparisons. All F statistics have 2 d.f. in the numerator.

Follow-up comparisons (Tukey's HSD) revealed that professional musicians scored higher than musically trained participants, who scored higher than the untrained group, on the MET Melody subtest, and on the Music Training, Perceptual Abilities, and Singing Abilities subscales from the Gold-MSI, *ps* < .005. This same pattern (i.e., professionals > trained > untrained) extended to the Active Engagement subscale and the General Factor from the Gold-MSI, *ps* < .001. The professional and trained groups scored higher than untrained participants on the MET Rhythm subtest and on the Emotions subscale from the Gold-MSI, *ps* < .001, but the professional and trained groups did not differ (Rhythm: *p* = .936, Emotions, *p* = .221). In short, expected group differences in musical ability were strong, whether performance was indexed objectively or by self-reports.

Table 8. *Descriptive and Inferential Statistics for the Personality Variables.*

	Musically trained		**Musically untrained**		**Professional musicians**		**Group comparison**		
	M	*SD*	*M*	*SD*	*M*	*SD*	$F_{(2,639)}$	*p*	η^2
Big Five									
Openness to experience	3.80	0.59	4.14	0.54	4.31	0.43	42.83	<.001	.118
Conscientiousness	3.50	0.69	3.63	0.68	3.80	0.71	4.23	.015	.012
Agreeableness	3.17	0.83	3.07	0.84	3.43	0.68	7.39	<.001	.022
Neuroticism	3.17	0.79	3.04	0.88	2.92	0.89	1.25	.287	.004
Metatraits									
Engagement	-.259	1.01	.023	.968	.492	.789	27.54	<.001	.078
Stability	-.183	.936	.101	.965	.290	1.07	6.83	.001	.019

Note. Age, education, and gender were held constant in the group comparisons. All F statistics have 2, 636 d.f.

For the Big Five traits, the three groups did not differ in terms of neuroticism, but they did on the other four traits. Descriptive and inferential statistics for personality variables are provided in Table 8. As expected, professional musicians and trained participants had higher mean openness scores compared to untrained participants, *ps* < .001, but the professional and trained groups did not differ, *p* = .132. Agreeableness showed a similar pattern, with professionals, *p* = .003, and trained participants, *p* = .013, scoring higher than nonmusicians, but no differences between the professional and trained groups, *p* = .984. Professional musicians had higher conscientiousness scores than untrained participants, *p* = .013, but the trained participants fell in between, such that they were no different from the professional, *p* = .604, or untrained, *p* = .296, groups. Finally, professional musicians were more extraverted than trained participants, *p* = .001, and untrained participants, *p* = .006, but the trained and untrained groups did not differ, *p* = .413.

For personality metatraits (see Table 8), engagement scores were higher for professional musicians compared to trained participants, *p* = .001, and untrained ones, *p* < .001, and higher for trained than for untrained participants, *p* = .019. Stability scores were higher for professional musicians, *p* = .004, and trained participants, *p* = .021, compared to untrained participants, but the professional and trained groups did not differ, *p* = .972.

Table 9. *Descriptive and Inferential Statistics for the Cognitive Variables.*

	Musically trained		**Musically untrained**		**Professional musicians**		**Group comparison**		
	M	*SD*	*M*	*SD*	*M*	*SD*	$F_{(2,639)}$	*p*	η^2
Mind wandering	3.29	0.96	3.13	0.88	2.88	0.95	2.75	.064	.008
Cognitive ability - 1	.614	.015	.654	0.15	.614	.014	3.59	.028	.010
Cognitive ability - 2	23.3	5.64	25.1	6.73	22.0	5.31	5.90	.003	.017

Note. Age, education, and gender were held constant in the group comparisons. All F statistics have 2, 636 d.f.

For cognitive variables, descriptive statistics and inferential statistics are provided in Table 9. The three groups did not differ in mind-wandering, but they did in general cognitive ability. As predicted, trained participants had higher scores than untrained participants, *p* = .048. Unexpectedly, trained participants also had higher scores than professional musicians, *p* = .035, who did not differ from untrained participants, *p* = .864. After adjusting for the covariates, professionals actually had the lowest mean. Because the professionals were older on average than the other groups, if their absolute (unadjusted) levels of performance matched that of the trained participants, this could potentially indicate higher-than-expected cognitive ability. Nevertheless, even when age was allowed to covary, the advantage remained evident for trained participants over professional musicians, *p* = .005, and untrained participants, *p* = .038, but the professional and untrained groups did not differ, *p* = .427.

We also considered whether the method of scoring the MaRs-IB played a role in response patterns, because it awarded the same score for (1) participants who took the maximum amount of time (30 s) for each item and were correct on 14 of 16 trials, and (2) those who completed 48 trials with 42 correct responses (i.e., proportion correct = .875 in both instances). Accordingly, we recalculated our measure of cognitive ability as the sum of correct responses, which is consistent with scoring of Raven's test, whether timed (Swaminathan et al., 2017) or untimed (Carpenter et al., 1990). Response patterns did not change. There was a main effect of group, with trained participants scoring higher than untrained participants, *p* = .012, and professional musicians, *p* =.003, who did not differ, *p* = .607. In absolute terms, mean scores (adjusted and unadjusted) were lowest for the professionals.

3.4 Discussion

In this study, we examined whether and how professional musicians and musically trained and untrained individuals differ in terms of musical ability, personality, and cognition. Compared to untrained participants, the musically trained and professional groups had higher scores on all measures of musical ability, the Big Five traits openness and agreeableness, and both personality metatraits. Being a professional musician was additionally predictive of even higher levels of musical ability, extraversion, and the metatrait engagement.

The musically trained group performed better than the untrained group on our test of cognitive ability, a finding that replicates previous results (for review see Swaminathan & Schellenberg, 2019). There was no evidence, however, of enhanced cognitive abilities among professional musicians, who scored significantly lower than trained participants, and no different from the untrained group. Since our sample size was very large, it is unlikely that statistical power played a role. It seems implausible, moreover, that professionals would exceed the trained participants in attempts to replicate our findings directly. This result is inconsistent with proposals that learning and performing music play a causal role in determining nonmusical cognitive abilities (e.g., Patel, 2011; Tierney & Kraus, 2013). Indeed, such hypotheses of far transfer and plasticity remain contentious (e.g., Degé, 2021; Sala & Gobet, 2020). As one example, Jäncke (2009) speculated that when a student learns to play a musical instrument, attention, planning functions, memory, and self-discipline are also practiced. It is therefore assumed that music experience would positively influence executive functions, language functions, and even intelligence in general (Jäncke, 2009). If this hypothesis were true, such effects might reach a plateau at some point, but they would be unlikely to go in reverse.

Nevertheless, our test of general cognitive ability was a single, brief test of abstract reasoning, even though general cognitive ability (*g*) is best measured as a latent variable extracted from a battery of tests that cover a wide range of abilities (Carroll, 1993). Clearly, a large battery of tests was unfeasible with our online testing context, such that our choice to administer the MaRs-IB was motivated primarily by practical reasons. As noted, however, the MaRs-IB is modeled after Raven's Advanced Progressive Matrices (Raven, 1965), which measures the ability to create abstract relationships and the ability to dynamically manage a wide range of problem-solving objectives within the working memory (Carpenter et al., 1990). Such abilities are considered central to virtually all concepts of intelligence, even those that attempt to expand its definition beyond "book smarts" (Sternberg, 1985). Although a clear limitation of the present study is that the results regarding cognitive ability could be test-specific, or specific to tests of matrix reasoning, our choice of test was justified by the testing context.

Our finding of elevated engagement and extraversion for professional musicians, but not for musically trained participants, seems intuitive because most professional musicians perform music publicly, at least at some point in their lives. Additionally, most of our professionals were music professors in addition to instrumentalists (≅ 72% in our sample), and education is in essence a social process. Engagement and extraversion have also been associated previously with creative behaviors, including music (Feist, 2019; Sylvia et al., 2009). Our results differ from those of Kuckelkorn et al. (2021), however, who documented high levels of extraversion among some subgroups of professional musicians (vocalists), but not others. In the current study, we found evidence of a more general effect, with group differences in engagement and extraversion being independent of instrument category. Our subgroups of participants per category were small (e.g., 16 professional vocalists), though, because we did not set out to explore instrument effects. Future research could explore the possibility of such effects in greater detail.

Although our results showed that professional musicians differ from other individuals primarily in terms of musical abilities and personality, there is no doubt that some musicians are very intelligent. IQ is also associated positively with eminence as a musician or composer, as it is across professions, although personality factors are as important as cognitive ability in predicting high levels of achievement (Miles, 1926; Simonton, 2006; 2009). The average professional musician, however, appears to differ from the general population primarily in terms of personality and musical ability rather than cognitive ability.

We propose that individual differences in musical ability, personality, and cognitive ability, in combination with contextual factors (e.g., socioeconomic status), jointly influence developmental trajectories of musical experience. Crucially, however, they contribute differently to predicting (1) who takes music lessons and for how long, and (2) who becomes a professional musician. During the childhood and teenage years, those who have high levels of musical ability, openness-to-experience, and cognitive ability, would tend to take music lessons for the longest duration (Corrigall et al., 2013; Kragness et al., 2021). Individuals with lower levels on one of these dimensions would be more likely to discontinue training or never begin, while those with lower levels on two (or three) dimensions would be even more likely to discontinue, probably at an earlier date. In early adulthood, most high-functioning individuals would opt to enter nonmusic professions because of personal interests, practical reasons (e.g., obtaining a well-paying job), or because of suboptimal levels of musical ability and/or personality characteristics. Other individuals, with high levels of musical ability and engagement (openness and extraversion), would be the most likely to choose a career in music. In some instances, individuals with high levels of musical ability, cognitive ability, and engagement might also pursue music further, or enter non-musical professions while maintaining their involvement in music. These proposals represent testable hypotheses that could be addressed in future developmental, longitudinal, and correlational studies.

Although our emphasis is on self-selection, which has typically been overlooked (Schellenberg, 2020), the environments people seek out undoubtedly influence who they become (Sauce & Matzel, 2018). In the case of skilled musical performance, the role of practice is incontrovertible. For objective measures of musical ability, however, music training plays a negligible role (Kragness et al., 2021). For cognitive ability and personality, shared environmental effects also appear to be small. Although the environment explains approximately half of the variance, these effects stem primarily from individual (nonshared) experiences (Harris, 2006).

4. Study 3: Musical ability: in-person and online testing[5]

4.1 Rationale and aims of the study

Internet is part of our life: more than half of the world's population now uses the internet, especially young people living in developed countries (98%; International Telecommunication Union, 2020). The COVID-19 pandemic increased the amount of time people spend on the internet to restrict in-person contact, making online testing an important option for psychological research. Online methods were used as an alternative to in-person research conducted in the laboratory even before the pandemic (e.g., Chetverikov & Upravitelev, 2015; Houben & Wiers, 2008; Milne et al., 2020; Taherbhai et al., 2012). Moreover, the development of a number of online platforms provided the possibility to recruit and test a considerable number of people (e.g., Gosling & Mason, 2015; Grootswagers, 2020).

Although there are some concerns about online testing, such as lack of control over characteristics of the samples and testing contexts (e.g., Birnbaum, 2004; Krantz & Dalal, 2000), online studies have several features that make them even better than in-person testing (e.g., Casler et al., 2013; Dandurand et al., 2008; Gosling et al., 2004): internet samples can be more representative of the general population (in terms of age, gender, and socioeconomic status); data quality can be similar; building online experiments, recruiting participants, and collecting data can be more efficient in terms of time and costs; access to specific target (e.g., musicians), is easier; participants may feel more comfortable at home than in a laboratory. On the other hand, despite these benefits, online testing needs careful experimental designs to maximize control (e.g., Gosling et al., 2004), and motivational strategies to engage people and improve the probability that they complete the whole experiment. Also, auditory and music research can be particularly challenging, since the contexts of

---

[5] This study has already been published: Correia, A. I., Vincenzi, M., Vanzella, P., Pinheiro, A. P., Lima, C. F., & Schellenberg, E. G. (2022). Can musical ability be tested online?. *Behavior research methods*, 54(2), 955-969. https://doi.org/10.3758/s13428-021-01641-2

online testing are more uncontrolled in terms of extraneous sounds, technical aspects of stimulus presentation, and potential interruptions (e.g., Milne et al., 2020). Although this variability can be reduced by asking participants to follow specific instructions (e.g., to wear headphones), experimental control remains limited.

Positive results regarding the possibility to obtain the same results from in-person and from online testing come from recent evidence. In particular, an online study about reinforcement learning (Nussenbaum et al., 2020), replicated a main effect of age that was reported in an earlier in-person study (Decker et al., 2016). In other developmental research, online data replicated a mediating role for abstract reasoning ability in the link between age and model-based learning (Chierchia et al., 2019).

An open question regarding the adaptability for online testing is whether longer and more cognitively demanding tasks, such as those measuring music abilities, can be similar to in-person testing.

In the present investigation, we used the platform Gorilla (http://www.gorilla.sc/; Anwyl-Irvine et al., 2020) to create an online version of an objective measure of musical ability—the Musical Ear Test (MET). The MET is a listening test that has documented reliability and validity (Swaminathan & Schellenberg, 2021; Wallentin et al., 2010). It is designed in the tradition of musical aptitude tests (see Chapter 1, paragraph 1.4), with two subtests, Melody and Rhythm, which require participants to determine whether two auditory sequences (a standard followed by a comparison) are identical or not, on multiple trials. Musical abilities' tests were designed to identify whether musically untrained individuals (especially children) were likely to benefit from music lessons. These tests, as well as more recent tests of musical ability (Fujii & Schlaug, 2013; Law & Zentner, 2012; Peretz et al., 2003, 2013; Ullén et al., 2014; Zentner & Strauss, 2017), all require comparisons between same or diverse auditory stimuli, differing in pitch or time, or along other dimensions (such as timbre or amplitude). These tests rely on core musical skills, specifically auditory short-term and working memory and perceptual discrimination. However, musical ability includes many other aspects of behavior (e.g., expert levels of performance) that are dependent on learning and practice; the goal of tests such as the MET is to measure musical ability in the absence of any formal training, and to do so objectively and quickly.

We also used Gorilla to run the entire testing session, which included measures of general cognitive ability and personality, and to create an online version of a self-report measure of musical behavior and expertise—the Goldsmiths Musical Sophistication Index (Gold-MSI; Lima et al., 2020; Müllensiefen, et al., 2014). The Gold-MSI served as our principal measure of construct validity. Virtually all developers of tests of musical ability report positive correlations with musical expertise as a means of documenting a test's validity (Asztalos & Csapó, 2014; Law & Zentner, 2012; Wallentin et al., 2010; Zentner & Strauss, 2017; Ullén et al., 2014). We compared response patterns from our online sample with previous studies that had large samples of participants: Swaminathan et al. (2021, N = 523) for the MET, and Lima et al. (2020, N = 408) for the Gold-MSI. We compared the present sample with the comparison samples in terms of psychometric characteristics (i.e., internal reliability, construct validity, correlations between subtests, and correlations between musical ability and musical sophistication).

Moreover, we tested for associations with demographic variables, cognitive ability, and personality, because previous studies have shown robust associations with these variables (e.g., Cooper, 2019; Greenberg et al., 2015; Kuckelkorn et al., 2021; Lima et al., 2020; Moreno et al., 2011; Swaminathan & Schellenberg, 2021). Absolute levels of performance on our measures could vary across samples depending on the degree to which they differ in music training, age, cognitive ability, personality, education, and so on. In terms of age and education, Lima et al. (2020) tested Portuguese individuals from the general population who varied widely, whereas Swaminathan et al. (2021) tested Canadian undergraduates who varied minimally.

The Gold-MSI has a history of online and in-person testing (Greenberg et al., 2015; Lima et al., 2020; Müllensiefen et al., 2014; Schaal et al., 2015); so, we predicted that results from our online version of the test would be similar to those from the paper-and-pencil administration of Lima et al.

(2020), with comparable psychometric properties. On the other hand, we were less certain of the outcome with the online version of the MET, considering that technological requirements were much greater for an objective listening test, which required participants to compare, on each of 52+52 trials, two auditory stimuli.

Therefore, the main aim of our study was to determine whether the MET could be successfully administered online. Evidence of success required that the test's internal reliability would not be compromised by online administration, that performance would correlate positively with musical expertise, and that musical ability would have positive associations with general cognitive ability. Moreover, as is the case with in-person testing (Swaminathan & Schellenberg, 2021), musical expertise should be a better predictor of scores for the Melody subtest of the MET than for the Rhythm subtest.

Previous evidence (e.g., Swaminathan & Schellenberg, 2019; Butkovic et al., 2015) indicated that the online test's success would be supported by a positive correlation with scores only one dimension from the Big Five model of personality (McCrae & Costa, 1987; McCrae & John, 1992), that is, openness to experience.

Another novel aspect of the present investigations included the prediction that mind-wandering would be negatively associated with performance on the MET, requiring 18 minutes concentration; one might expect lower mind-wandering levels among individuals who have taken music lessons for longer time, since learning to play music requires much effort and focus.

The Gold-MSI allowed us to explore whether aspects of musical expertise other than training were predictive of performance, and whether their predictive power would vary across subtests. Moreover, it allowed to examine whether musical ability would be associated with active engagement with music, emotional responding to music, and self-reports of singing and perceptual abilities; such associations would confirm that the narrow range of abilities tested by the MET is predictive of a much broader range of musical abilities.

4.2 Method

4.2.1 Participants

The original sample included a total of 754 participants. Persons who did not complete the MET (n = 100), or failed to respond on more than 10 trials in total (n=39) or more than 5 in a row (n= 7) (on either the Melody or the Rhythm subtest), were subsequently excluded. The final sample included 608 participants (361 female, 243 male, 4 unreported) between 18 and 88 years of age (M = 34.2, SD = 15.1). Most of them completed high school (n = 207), or had a university degree (Bachelor's, n = 108, Master's, n = 191, PhDs, n = 58). Education data were missing for 41 participants.

Participants were recruited primarily through social media posts and snowball sampling. The experiment was available in four languages, and participants had to complete it in their native language (Italian, n = 288; European Portuguese, n = 153; Brazilian Portuguese, n = 123; English, n = 44). Some of the participants had no history of music lessons (n = 151) or a maximum of 2 years (n = 133), 156 had 10 years or more. The training included private lessons (n = 123), or classes taught at university (n = 122) or in musical academies or conservatories (n = 84). Others (n = 85) were self-taught. On average, participants with music lessons started their training at the age of 11.4 years (SD = 7.1; range: 2–56).

4.2.1 Materials

All questionnaires and tasks were created originally in English. They were then adapted for online testing through Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). Validated translations of the measures were used when available (e.g., the Big Five Inventory in European-Portuguese and Italian). Otherwise, translations by bilinguals who were native speakers and also fluent in English were created and proposed.

Objective behavioral tests

Musical ability

We created an online version of the *Musical Ear Test* (MET; Wallentin et al., 2010) to evaluate music perception abilities. As in the original version, the online MET had two subtests, Melody (52 trials) and Rhythm (52 trials). On each trial, participants listened to two short musical excerpts, and they had to make a Yes/No judgment about whether the stimulus was the same as the standard. Half of the trials were same, and half were different, for both Melody and Rhythm subtests; also, the stimuli and order of presentation were the same as in the original lab test. All musical stimuli had the same structure (4/4 time) and tempo (100 beats per minute). A lower-amplitude metronome sound indicated the underlying beat. Two practice trials, providing feedback, preceded each subtest (one same, one different), while feedback was not provided for test trials.

In the original test, instructions and trials are presented through an audio file, and a male speaker provides task instructions and the number of each trial. Trials are not self-paced: in the original in-person version, participants are given a response sheet, and they have a brief window after each trial (1500 ms for melodic trials, 1659 to 3230 ms for rhythmic trials) to respond by checking yes or no. In our online, instructions and trial numbers were converted to text, and participants had to read them. The stimuli from each trial were digitally copied from the original audio file and the duration of the inter-stimulus intervals was preserved, for a total duration of approximately 20 min, identical to the in-person version. The trial number and the question asking whether the melodic/rhythmic phrases were identical were visible on the screen from the beginning of each trial until the participant responded. Immediately after the audio stimulus ended, two Yes / No buttons appeared, and participants had a few moments to respond by clicking the appropriate button, before the next trial started automatically. To enhance the online testing experience, a progress bar at the bottom of the screen was also provided throughout both Melody and Rhythm subtests, to give participants the opportunity to monitor where they were in relation to the beginning and end of each subtest. At the end of the test, feedback about the participant's performance was provided. It was calculated as the total number of correct responses on the Melody and Rhythm subtests. For statistical analyses, a Total score was also calculated as the sum.

General cognitive ability

The Matrix Reasoning Item Bank (MaRs-IB; Chierchia et al., 2019) was our measure of general cognitive ability. It is an online test of abstract (nonverbal) reasoning modeled after Raven's Advanced Progressive Matrices (Raven, 1965), consisting of 80 trials, in which a 3 × 3 matrix is present on the computer screen. Only bottom-right cell (ninth) was always empty, while the other eight cells contained abstract shapes: participants had to complete the matrix by choosing one of four alternatives. Associations among shapes could vary on a single dimension (e.g., color), up to four dimensions (e.g., color, size, shape, and location). Before the matrix was presented, a 500-ms fixation cross appeared in the middle of the screen, followed by a 100-ms white screen, on each trial. Participants then had up to 30 s to select their response: if participants responded, the trial ended earlier; if no response was provided after 25 s, a clock appeared and indicated the time remaining. The order of the trials was the same for all participants. The first five items served to familiarize participants with the task, and were easier. The duration of the entire task was fixed at 8 minutes; however, participants were not informed of the task duration or the number of trials, and if they completed the 80 trials in less than 8 minutes, the trials were presented again following the same order, but responses from the second round were not considered in calculating scores. Scores were calculated as the proportion of the total number of responses given by the participant that were correct. For the statistical analyses, proportions were logit-transformed.

Questionnaires

Musical expertise

The *Gold-Musical Sophistication Index* was our measure for tests of construct validity (Müllensiefen et al., 2014). It is a self-report questionnaire of musical expertise and behavior, consisting in 38-items, evaluating different aspects of musical behaviors and abilities. Scores on different subsets of items are averaged to form five subscales: Active Engagement (e.g., I’m intrigued by musical styles I'm not familiar with and want to find out more), Perceptual Abilities (e.g., I can tell when people sing or play out of tune), Music Training (e.g., reversed: I have never been complimented for my talents as a musical performer), Singing Abilities (e.g., I am able to hit the right notes when I sing along with a recording), and Emotions (e.g., I often pick certain music to motivate or excite me). A General Musical Sophistication factor is also formed, averaged from 18 items representative of the five subscales. Each item is provided with a 7-point rating scale.

Personality

Personality traits were evaluated with the *Big Five Inventory* (BFI, John & Srivastava, 1999), a self-report questionnaire with 44 items that assess five dimensions of personality: openness to experience (10 items), conscientiousness (9 items), extroversion (8 items), agreeableness (9 items), and neuroticism (8 items). Items are mixed in terms of presentation order. Participants rated how much each expression describes them using a five-point rating scale (1 = disagree strongly, 5 = agree strongly). The BFI was initially published in English (John & Srivastava, 1999), and then translated into European-Portuguese (Brito-Costa et al., 2015) and Italian (Ubbiali et al., 2013). A Brazilian-Portuguese version was created by modifying the European-Portuguese version, double- checking the original English version for fidelity.

Mind-wandering

A measure of sustained attention and ability to focus was the Mind-Wandering Questionnaire (MWQ, Mrazek et al., 2013), a five-item scale with good psychometric properties that evaluates trait levels of mind-wandering (e.g., I find myself listening with one ear, thinking about something else at the same time). Participants rated how much they agreed with each sentence on 1 (almost never) to 6 (almost always) scale.

4.2.2 Procedure

All questionnaires and tasks were completed in a single testing session. A hyperlink posted on social media (e.g., Facebook, Twitter, LinkedIn) provided the initial access to the experiment. A brief description of the study, specifying that participants should complete the testing session in a quiet room with a stable internet connection, use headphones, and turn off sound notifications from other devices and applications (e.g., email, phone messages), was provided, together with information regarding the duration of the experiment of approximately 40 min.

Informed consent and some basic demographic questions (e.g., age, gender, education) were proposed at the beginning. Then, participants completed the self-report questionnaires in a fixed order (MWQ, Gold-MSI, and BFI). After, they were tested on the MaRs-IB and finally the MET. At the end of the experiment, participants received a brief feedback about their scores on personality, musical sophistication, and musical ability measures. A final open-ended box asked participants to describe any problems that might have occurred during the testing session.

4.3 Results

The statistical analyses incorporated standard frequentist null-hypothesis testing, as well as Bayesian analyses conducted with JASP version 0.14.1 (JASP Team, 2020) using default priors[6], as in the reports from the comparison samples (see below, Lima et al., 2020; Swaminathan et al., 2021). Considering the large sample, very small effects were statistically significant with null-hypothesis testing. For example, with N = 608, correlations greater than .08 in absolute value were significant with $p < .05$. We considered small associations to be reliable only if they also passed a conventional threshold for what is considered substantial evidence using Bayesian statistics (Jarosz & Wiley, 2014; Jeffreys, 1961). Specifically, when the Bayes factor (BF_{10}) was greater than 3.00, the observed data

---

[6] Correlations, stretched beta prior width = 1; *t* tests, zero-centered Cauchy prior with scale parameter 0.707; linear regressions, JZS prior of $r = .354$; Wagenmakers et al., 2018; Wagenmakers et al., 2016).

were at least three times as likely under the alternative as the null hypothesis. Lower values ($1.00 < BF_{10} < 3.00$) indicated that the data provided evidence for the alternative hypothesis that was considered to be weak or anecdotal. If $BF_{10} < 1.00$, the observed data provided evidence that favored the null hypothesis in a reciprocal manner (i.e., substantial evidence when $BF_{10} < .333$). More extreme values provided strong ($BF_{10} > 10.0$ or $< .100$), very strong ($BF_{10} > 30.0$ or $< .033$), and decisive ($BF_{10} > 100.0$ or $< .010$) evidence for either the alternative or null hypothesis, respectively.

Initial analyses examined whether the present sample differed from comparison samples (Lima et al., 2020; Swaminathan et al., 2021) in terms demographic variables (i.e., gender, age, years of education, and of musical training).

Demographic variables

As for gender, although women were the majority in all three samples, the proportion who were men was decisively higher in the present sample (40.2%) than in Lima et al. (25.0%), $\chi^2(1, N = 1012) = 25.14$, $p < .001$, $BF_{10} > 100$. It was also higher than in Swaminathan et al. (32.4%), $\chi^2(1, N = 1122) = 7.31$, $p = .007$, $BF_{10} = 2.99$, but the observed data provided only weak evidence for a group difference.

Mean age of participants did not differ between the present sample ($M = 34.22$, $SD = 15.11$) and that of Lima et al. ($M = 32.95$, $SD = 14.38$), $p = .181$, $BF_{10} = 0.17$, with substantial evidence favoring the null hypothesis. The present sample was considerably older than the undergraduates tested by Swaminathan et al. ($M = 19.04$, $SD = 2.03$), $t(632.36) = 24.52$, $p < .001$ (unequal variances test), and the variance was greater, $F(607, 522) = 892.49$, $p < .001$. The observed data provided decisive evidence for the group difference in age, $BF_{10} > 100$.

In terms of years of education, the sample of undergraduates from Swaminathan et al. (2021) varied minimally. Compared to the sample from Lima et al. (2020, $M = 6.94$, $SD = 2.11$), the present sample had less education ($M = 6.16$, $SD = 1.05$), $t(552.19) = 6.84$, $p < .001$ (unequal variances test), Cohen's $d = .490$, $BF_{10} > 100$, and less variance in terms of age, $F(565, 407) = 326.67$, $p < .001$.

Regarding musical training, participants in the present sample had a higher level ($M = 4.26$, $SD = 2.30$) than those in the sample from Lima et al. ($M = 2.68$, $SD = 1.93$), $t(966.32) = 11.84$, $p < .001$ (unequal variances test), Cohen's $d = .744$, $BF_{10} > 100$, and more variability in training, $F(607, 407) = 20.98$, $p < .001$. Since Swaminathan et al. (2021) treated duration of music training as a continuous variable, we re-coded the variable so that it conformed to item 36 from the Gold-MSI (ordinal scale), thereby making duration of training comparable across samples. The present sample had more music training than participants tested by Swaminathan et al. ($M = 3.78$, $SD = 2.40$), $t(1089.60) = 3.42$, $p < .001$ (unequal variances test), Cohen's $d = .204$, $BF_{10} = 21.4$, although variance was greater in the previous sample, $F(607, 522) = 7.10$, $p = .008$.

To sum up, the present sample had a larger proportion of participants who were men, and the mean age was higher than in Swaminathan et al. (2021) but similar to Lima et al. (2020), and mean levels of music training were higher in the present sample than in both comparison samples.

Personality

In Swaminathan et al. (2021) personality data are not reported. Comparisons of the present sample with Lima et al. (2020)'s are provided in Table 10.

The present sample had higher levels of openness-to-experience and neuroticism, and lower levels on conscientiousness, extroversion, and agreeableness. The observed data provided decisive evidence for group differences in openness-to-experience and agreeableness, very strong evidence for differences in extroversion, and substantial evidence for differences in conscientiousness and neuroticism. Variance in agreeableness was also notably greater for the previous sample, $F_{(607, 394)} = 7.75$, $p = .005$.

Table 10. *Comparisons of personality data from the BFI between the present sample and Lima et al. (2020).*

	Present Sample *N* = 608 Mean (*SD*)	Lima et al. (2020) *N* = 395 Mean (*SD*)	*p*	Cohen's *d*	BF_{10}
Openness to experience	4.01 (.58)	3.57 (.66)	< .001	.721	> 100
Conscientiousness	3.58 (.70)	3.70 (.64)	.003	.190	4.92
Extroversion	3.19 (.79)	3.38 (.78)	< .001	.241	65.4
Agreeableness	3.79 (.54)	3.95 (.48)	< .001	.306	> 100
Neuroticism	3.11 (.82)	2.95 (.84)	.002	.201	8.19

Correlations Between Predictor Variables and the Gold-MSI

The main analyses focused on musical ability, musical experience, and their correlates, including demographics (age, gender, education), cognitive ability, personality, and mind-wandering. Pairwise correlations among potential predictors are provided in Table 11. Multiple associations were evident, many for which the observed data provided decisive evidence. With increasing age, participants were more likely to be men, to have more education, and to score higher on openness-to-experience, conscientiousness, and extroversion. Older individuals also tended to have lower scores on our measures of cognitive ability, mind-wandering, and neuroticism. Analyses of gender revealed that women scored higher than men on two personality dimensions: agreeableness and neuroticism. Amount of education was correlated positively with cognitive ability and conscientiousness, but negatively with mind-wandering and neuroticism. Cognitive ability had no additional associations with mind-wandering or personality. Mind-wandering was associated negatively, however, with conscientiousness, extroversion, and agreeableness, but positively with neuroticism. Associations among personality variables also revealed considerable overlap. Extroversion was correlated positively with openness-to-experience, conscientiousness, and agreeableness, but negatively with neuroticism; neuroticism had additional negative associations with conscientiousness and agreeableness; and conscientiousness was correlated positively with agreeableness.

Table 11. *Pairwise associations (Pearson correlations and Bayes factors) among potential correlates of musical sophistication and musical ability.*

		Gender	Education	Cognitive Ability	Mind-Wandering	Openness	Conscien-tiousness	Extro-version	Agree-ableness	Neuro-ticism
Age	*r*	.204	.450	-.177	-.378	.143	.276	.180	.094	-.276
	BF_{10}	>100	>100	>100	>100	26.4	>100	>100	.753	>100
Gender	*r*		.095	-.023	-.080	.074	-.029	-.046	-.126	-.207
	BF_{10}		.688	.060	.346	.263	.065	.096	6.62	>100
Education	*r*			.125	-.261	.118	.263	.096	.062	-.170
	BF_{10}			4.41	>100	2.72	>100	.725	.155	>100
Cognitive Ability	*r*				.097	.010	-.055	-.025	-.060	-.023
	BF_{10}				.862	.052	.129	.061	.152	.060
Mind-Wandering	*r*					-.046	-.626	-.174	-.241	.409
	BF_{10}					.096	>100	>100	>100	>100
Openness	*r*						.083	.262	.096	-.070
	BF_{10}						.409	>100	.857	.223
Conscientiousness	*r*							.221	.201	-.359
	BF_{10}							>100	>100	>100
Extroversion	*r*								.135	-.209
	BF_{10}								13.6	>100
Agreeableness	*r*									-.314
	BF_{10}									>100

We had no hypotheses about the testing language of the online study, and exploratory analyses confirmed that musical ability did not vary as a function of language when individual differences in age, education, cognitive ability, and openness to experience were held constant. In fact, for the Melody subtest, the Rhythm subtest, and Total scores of the MET, the observed data provided substantial evidence for the null hypothesis (all $\text{BF}_{10} < .250$). Testing language was not considered further.

Musical expertise

Because of the large number of musicians in the current sample, mean scores were higher than they were in Lima et al. across Gold-MSI Subtests and the General Factor, *ps* < .001, all $\text{BF}_{10} > 100$ (Table 12).

Table 12. *Reliability (Cronbach's alpha) and descriptive statistics for the Gold-MSI Subtests and the General Factor. Data are presented for the whole sample and separately for the unpublished Italian and Brazilian versions. For comparison purposes, values from Lima et al. (2020) are also provided.*

	Current Online Sample											
	Whole sample (*N* = 608)			Italian Gold-MSI (*n* = 288)			Brazilian Gold-MSI (*n* = 123)			Lima et al., 2020 (*N* = 408)		
Gold-MSI Subtest	α	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*
Active Engagement	.85	4.36	1.21	.88	4.17	1.29	.83	4.58	1.20	.85	3.67	1.15
Perceptual Abilities	.87	5.53	1.07	.88	5.46	1.15	.86	5.67	1.04	.85	4.95	0.97
Music Training	.92	3.96	1.83	.93	3.92	1.96	.91	4.45	1.70	.89	2.64	1.45
Singing Abilities	.84	4.42	1.36	.87	4.34	1.48	.78	4.57	1.23	.83	3.81	1.21
Emotions	.69	5.75	0.86	.68	5.63	0.85	.74	5.83	0.96	.82	5.22	1.06
General Factor	.92	4.41	1.29	.94	4.27	1.42	.90	4.73	1.14	.91	3.60	1.07

As in the comparison sample and in Müllensiefen et al., (2014), pairwise correlations among Gold-MSI scores were all positive, and the observed data provided decisive evidence for an association in each instance (Table 13).

Table 13. *Pairwise Correlations Among the Gold-MSI Subtests and the General Factor (N = 608).*

	Perceptual Abilities	Music Training	Singing Abilities	Emotions	General Factor
Active Engagement	.562	.526	.553	.600	.735
Perceptual Abilities		.667	.740	.534	.826
Music Training			.632	.337	.863
Singing Abilities				.496	.874
Emotions					.548

Note. All p-values < .001. All $BF_{10} > 100$.

Examination of correlations between Gold-MSI scores and potential predictor variables revealed a relatively small number of instances in which the observed data provided substantial or stronger evidence for an association (Table 14).

Table 14. *Associations (Pearson Correlations and Bayes Factors) Between Scores on the Gold-MSI and Demographic Variables, Cognitive Ability, Mind Wandering, and Personality.*

		Active Engagement	Perceptual Abilities	Music Training	Singing Abilities	Emotions	General Factor
Age	*r*	.009	.088	.113	.029	-.162	.078
	BF_{10}	.052	.524	2.46	.066	>100	.314
Gender	*r*	.115	.094	.132	.025	-.050	.124
	BF_{10}	2.78	.734	10.3	.061	.108	5.30
Education	*r*	-.090	.049	.079	.014	-.088	.004
	BF_{10}	.513	.104	.306	.056	.474	.053
Cognitive Ability	*r*	-.062	.036	.008	.013	-.013	-.008
	BF_{10}	.164	.075	.052	.053	.054	.052
Mind-Wandering	*r*	-.062	-.099	-.135	-.072	.073	-.096
	BF_{10}	.160	1.01	13.6	.239	.254	.854
Openness	*r*	.469	.411	.405	.396	.403	.481
	BF_{10}	>100	>100	>100	>100	>100	>100
Conscientiousness	*r*	.040	.095	.084	.083	-.010	.071
	BF_{10}	.082	.786	.426	.420	.052	.230
Extroversion	*r*	.052	.113	.055	.192	.094	.119
	BF_{10}	.114	2.40	.126	>100	.744	3.88
Agreeableness	*r*	.101	.081	.127	.101	.148	.127
	BF_{10}	1.12	.362	6.96	1.11	40.6	6.83
Neuroticism	*r*	.011	-.030	-.081	-.020	.109	-.040
	BF_{10}	.053	.067	.364	.057	1.82	.082

Note. Gender was dummy-coded (Females = 0, Males = 1)

Evidence of a negative association between age and scores on the Emotions subtest emerged for demographic variables (age, gender, education). Also, strong evidence that men had more Music Training than women emerged, and there was substantial evidence for a male advantage on the General Factor. Cognitive ability had no significant associations with Gold-MSI scores, and the observed data provided substantial (or strong) evidence for the null hypothesis for all subtests. There was strong evidence for a small, negative association between mind-wandering and the Music Training subtest, but mind-wandering was not associated with any other Gold-MSI score, as expected.

Openness to experience personality trait was positively associated with all Gold- MSI scores (*rs* ≥ .4). Decisive and substantial evidence for positive but small associations between extroversion and Singing Abilities, and between agreeableness and Music Training, respectively (*rs* ≤ .2) emerged.

Musical ability

Statistics from tests of internal reliability for the online MET are provided in Table 15. Cronbach's alphas were virtually identical to those reported by the Wallentin et al., 2010, and higher than those reported in Swaminathan et al., (2021) comparison sample. Split-half (odd–even) reliabilities with Spearman-Brown formula were also higher than those reported by Swaminathan et al. (2021). To sum up, the internal reliability of the MET was not compromised by the online testing format.

Table 15. *Reliability statistics, including Cronbach's alpha and split-half (odd-even) correlations (Spearman-Brown formula), for scores on the MET. For comparison purposes, values from two previous reports are provided*

		Melody	Rhythm	Total
Current online sample (N=608)				
	Cronbach's alpha	.82	.70	.85
	Split-half correlation	.84	.75	.87
Swaminathan et al., 2021 (N=523)				
	Cronbach's alpha	.73	.62	.78
	Split-half correlation	.71	.68	.78
Wallentin et al., 2010, (N=69)				
	Cronbach's alpha	.82	.69	.85

Descriptive statistics for the Melody, Rhythm, and Total scores are provided in Table 16. For the entire sample, the observed means were higher than those reported by Swaminathan et al. (2021) for the Melody, Rhythm, and Total scores, as confirmed by independent-samples *t* tests, *ts*(1129) = 5.06, 5.90, and 6.23, respectively, *ps* < .001, all BF_{10} > 100. These findings were not meaningful, however, considering the musicianship's sample differences: to rectify this problem, we considered individuals with no music training (see Table 15). For participants with no music training, mean performance did not differ from that reported previously on the Melody subtest, *p* = .202, BF_{10} = .263, the Rhythm sub- test, *p* = .053, BF_{10} = .725, or for Total scores, *p* = .064, BF_{10} = .625, although evidence favoring the null hypothesis was substantial only for the Melody subtest. However, online-generated scores were comparable to in-person scores when they were expected to be comparable.

Table 16. *Descriptive statistics for scores on the MET. Melody and Rhythm scores were calculated from 52 trials. Total calculated from 104 trials. For comparison purposes, Swaminathan et al. (2021) are provided.*

	Current online sample			Swaminathan et al., 2021		
Whole sample						
	N	*M*	*SD*	*N*	*M*	*SD*
Melody	608	37.88	6.60	523	36.05	5.26
Rhythm	608	38.29	5.25	523	37.47	4.49
Total	608	76.17	10.54	523	72.52	8.89
No music training						
	N	M	SD	N	M	SD
Melody	151	34.91	6.44	189	34.15	4.41
Rhythm	151	36.66	5.79	189	35.56	4.68
Total	151	71.56	10.90	189	69.71	7.48

Melody and Rhythm scores were positively and decisively correlated, $r = .551$, N = 608, $p < .001$, $BF_{10} > 100$, with the magnitude of the association no different from that reported by Swaminathan et al. (2021), $r = .489$, $p = .154$, and Wallentin et al. (2010), $r = .520$, $p = .754$[7]. The data provided substantial evidence that performance did not differ between subtests, $BF_{10} = .214$, as in the earlier reports.

Demographics, cognitive ability, mind-wandering, and personality: correlations

Correlations between MET scores and demographic variables, cognitive ability, mind-wandering, and personality are provided in Table 17.

The observed data provided decisive evidence that as listeners increased in age, education, or cognitive ability, performance on the MET tended to improve as well. The one exception was the association between cognitive ability and Melody scores, for which the data provided substantial rather than decisive evidence. The correlation with cognitive ability was also higher for the Rhythm than for the Melody subtest, $z = 2.87$, $p = .004$.

For mind-wandering, there was substantial evidence for a negative association with scores on the Melody subtest, but no evidence of an association with Rhythm or Total scores. Nevertheless, the magnitude of the association was not significantly stronger for Melody than for Rhythm, $p > .1$. For personality, the observed data provided decisive evidence for positive associations between openness to experience and MET performance, but no evidence for associations with any other personality variable. In fact, all Bayes factors were below 1 with a single exception, and for two personality traits (conscientiousness, extroversion), the observed data provided substantial evidence for the null hypothesis.

---

[7] Comparisons of the magnitude of correlations were conducted with *Psychometrica* (https://www.psychometrica.de/correlation.html).

Table 17. *Pairwise associations (Pearson correlations and Bayes factors) between scores on the MET and demographic variables, cognitive ability, mind-wandering, and personality.*

		Melody	**Rhythm**	**Total**
Age	*r*	.206	.167	.214
	BF_{10}	>100	>100	>100
Gender	*r*	.099	.029	.077
	BF_{10}	1.01	0.066	0.306
Education	*r*	.209	.200	.232
	BF_{10}	>100	>100	>100
Cognitive ability	*r*	.131	.239	.204
	BF_{10}	9.84	>100	>100
Mind wandering	*r*	-0.122	-0.060	-0.107
	BF_{10}	4.60	0.153	1.63
Openness to experience	*r*	.241	.182	.243
	BF_{10}	>100	>100	>100
Conscientiousness	*r*	.068	.030	.058
	BF_{10}	0.210	0.067	0.142
Extroversion	*r*	.065	.069	.076
	BF_{10}	0.180	0.218	0.288
Agreeableness	*r*	.092	.060	.088
	BF_{10}	0.650	0.151	0.527
Neuroticism	*r*	-0.101	-0.018	-0.072
	BF_{10}	1.11	0.056	0.245

Note. Gender was dummy-coded (female = 0, male = 1). Ns = 608 except for education, n = 566.

Musical expertise and music training: correlations

Our main tests of construct validity involved correlations between scores on the MET and those from the subtests and General Factor from the Gold-MSI, which are provided in Table 18.

Table 18. *Pairwise associations (Pearson correlations and Bayes factors) between scores on the MET and scores on the Gold-MSI (N = 608).*

		Melody	Rhythm	Total
Active Engagement	*r*	.303	.186	.284
	BF_{10}	>100	>100	>100
Perceptual Abilities	*r*	.459	.320	.450
	BF_{10}	>100	>100	>100
Music Training	*r*	.491	.296	.458
	BF_{10}	>100	>100	>100
Singing Abilities	*r*	.406	.259	.386
	BF_{10}	>100	>100	>100
Emotions	*r*	.206	.141	.201
	BF_{10}	>100	22.5	>100
General Factor	*r*	.504	.307	.471
	BF_{10}	>100	>100	>100

All correlations were positive and statistically significant ($p < .001$), with the observed data providing decisive evidence for an association in each instance, except for the association between the Emotions subtest and Rhythm scores, which was strong but not decisive.

In Swaminathan et al., (2021) comparison sample, music training proved to be a better predictor of Melody than of Rhythm scores. The Gold-MSI scores showed a similar pattern. For Perceptual Abilities, Music Training, Singing Abilities, and the General Factor, correlations with the Melody subtest were higher than those for the Rhythm subtest, $zs > 4$, $ps < .001$. The same finding was weaker yet still evident for Active Engagement, $z = 3.16$, $p = .002$, but not for the Emotions subtest, $p = .086$.

Music Training subtest was the focus of additional analyses: associations between Music Training and MET scores were higher than those in the comparison sample (Swaminathan et al.,

2021) (see Table 18). This could be due to differences in how training was measured and/or a consequence of greater variability due to the higher proportion of musicians in the present sample.

The correlations were lower than correlations between MET scores and current daily practice reported by Wallentin et al. (2010), a likely consequence of differences in measurement.

Also, we asked whether performance on the MET was associated with the age at which music training began. As in Swaminathan et al. (2021), we considered only participants who had any training (n = 415) and divided them into two groups: those who started by age 7 (*early starters*) (n = 120) and, those who started at an older age (*late starters*) (n = 295). This split was theoretically motivated, based on the proposal of a sensitive period that extends up to 7 years of age, during which plasticity is greater and music training is presumed to have a stronger impact on development (Penhune, 2019, 2020; Penhune & De Villiers-Sidani, 2014). The results were similar to those reported in the comparison sample (Swaminathan et al., 2021): *early starters* had higher scores than *late starters* on the Melody subtest, $t(413) = 3.18$, $p = .002$, $BF_{10} = 14.7$, and on Total scores, $t(413) = 2.96$, $p = .003$, $BF_{10} = 7.82$, but not on the Rhythm subtest, $p = .076$, $BF_{10} = .543$. Nevertheless, early starters also had more Music Training, $t(413) = 4.11$, $p < .001$, $BF_{10} > 100$. When Music Training was held constant, the advantage for early starters disappeared for the Melody subtest, $p = .078$, $BF_{10} =.577$, and for Total scores, $p = .083$, $BF_{10} = .527$, although the observed data did not provide strong evidence for the null hypothesis.

Multiple regression analysis

We used multiple regression to determine which correlates made independent contributions in predicting performance on the MET in a final set of analyses. Specifically, we modeled MET Melody, Rhythm, and Total scores from a linear combination of variables, each of which had a reliable simple association with MET scores: age, education, cognitive ability, mind-wandering, openness to experience, and the Gold-MSI subtests. The results are summarized in Table 19.

Table 19. *Multiple regression results predicting MET scores from age, education, openness to experience, cognitive ability, mind-wandering, and the five Gold-MSI subtests.*

	Melody			**Rhythm**			**Total**		
Model									
R^2	.332			.210			.335		
Adjusted R^2	.320			.196			.323		
$F(10,555)$	27.63			14.76			27.98		
p	<0.001			<0.001			<0.001		
Predictors									
	β	*p*	BF_{10}	β	*p*	BF_{10}	β	*p*	BF_{10}
Age	.154	<0.001	610.7	.159	<0.001	>100	.177	<0.001	>100
Education	.098	.016	1.43	.089	.045	0.760	.107	.009	2.14
Cognitive Ability	.145	<0.001	>100	.259	<0.001	>100	.222	<0.001	>100
Mind-wandering	.010	.802	0.129	.013	.751	0.146	.013	.739	0.131
Openness to experience	-0.027	.523	0.129	-0.005	.918	0.160	-0.019	.648	0.125
Active engagement	.049	.347	0.218	.077	.174	0.306	.070	.178	0.357
Perceptual abilities	.177	.003	>100	.174	.008	>100	.199	<0.001	>100
Music training	.305	<0.001	>100	.128	.019	6.62	.256	<0.001	>100
Singing abilities	.053	.337	0.232	-0.019	.749	0.146	.023	.673	0.155
Emotions	-0.002	.972	0.140	-0.006	.908	0.169	-0.004	.931	0.148

For the Melody subtest, the Rhythm subtest, and Total scores, the overall model was significant, with independent and positive partial associations with age, education, cognitive ability, and the Perceptual Abilities and Music Training subtests from the Gold-MSI.

In the Bayesian counterpart to multiple regression, we first identified which model was most likely given the observed data. For the Melody subtest and for Total scores, it included age, education, cognitive ability, Perceptual Abilities, and Music Training – a finding that corroborated the frequentist results. We calculated a Bayes factor for each predictor by removing them from the model one at a time. As shown in Table 19, the observed data provided decisive evidence for the inclusion of Perceptual Abilities and Music Training in the model, and very strong (Melody) or decisive (Total) evidence for including cognitive ability and age. For education, however, the Bayes factor was less than 3. We calculated BF_{10} for the other (excluded) five variables by adding each to the model one at a time. For each variable, the observed data provided substantial evidence for the null hypothesis. In other words, the observed data were more likely with a model that did not include these variables.

For the Rhythm subtest, the best model of the data included age, cognitive ability, Perceptual Abilities, and Music Training. The observed data provided decisive evidence for the inclusion of age, cognitive ability, and Perceptual Abilities in the model, but only substantial evidence for including Music Training. For the other six variables, the observed data provide substantial evidence for the null hypothesis with one exception: they were more or less equally likely with a model that included or excluded education.

4.4 Discussion

The main aim if this study was to determine if an established and validated test of musical ability could be administered successfully on-line.

It is worth saying that approximately 20% of the sample who started the testing session did not complete it or provide usable data; however, this level of attrition is not surprising, considering that there was no compensation for participants to complete the session, other than receiving final feedback about personality, musical expertise, and musical ability. Also, the testing session was relatively long and, unlike in a laboratory, there were no research assistants motivating participants to continue the test. However, the findings were otherwise unequivocally positive: the results for the MET were both novel and noteworthy, since this objective listening test of musical ability was not adapted previously for online testing.

The main variable for testing construct validity was the Gold-MSI, and it was a proof of concept that the present sample of online participants would respond similarly to a sample of participants tested in paper and pencil format (Lima et al., 2020). Indeed, response patterns to the online Gold-MSI were very similar to those reported previously. The internal reliability of the test was in fact similar, except for the Emotions subtest. As in the earlier study, age correlated negatively with the Emotions subtest, even though Lima et al. reported a negative correlation between age and *all* Gold-MSI subtests. These discrepancies in response patterns between samples could stem from differences in music training: compared to the previous study, our subsample of participants with very high levels of music education was higher. In fact, 25.6% of our sample had 10 or more years of music lessons, compared to the 5.6% in Lima et al.'s. Since increases in musical experience are accompanied by increases in age, a negative association between age and Gold-MSI scores would be less likely in our online sample. Despite these differences in samples, correlations among Gold-MSI subtests, and between Gold-MSI scores and personality variables were similar across testing formats. Instead, there was little evidence of an association between cognitive ability and the Music Training subtest from the Gold-MSI. While in childhood music training is often correlated positively with cognitive ability (Corrigall et al., 2013; Corrigall & Schellenberg, 2015; Kragness et al., 2021; Schellenberg, 2006, 2011; Schellenberg & Mankarious, 2012; Swaminathan & Schellenberg, 2020), in adulthood such associations tend to be weaker (Lima & Castro, 2011; Schellenberg, 2006). When matrix-type tests of cognitive ability, such as Raven's test and the test used in the present sample (MaRs-IB), are given to students from an introductory psychology course, positive associations with music training are evident in some instances (Swaminathan et al., 2017, 2018, 2021) but not in others (Schellenberg & Moreno, 2010; Swaminathan & Schellenberg, 2017). These associations may become less likely in samples of older participants with a large proportion of professional musicians (Lima & Castro, 2011).

The internal reliability of the online version of the MET proved to be similar to in-person administration (Wallentin et al., 2010; Swaminathan et al., 2021). Our results confirmed that the correlation between Melody and Rhythm subtests did not differ across formats. Additionally, there was no difference in performance between subtests. Also, when the present and comparison samples were equated for music training by focusing solely on participants with no training, average levels of performance were similar. Moreover, there were no gender differences in performance on the MET, as in the comparison sample. Finally, performance was strongly associated with openness to experience, but not with other dimensions of personality (Greenberg et al., 2015; McCrae & Greenberg, 2014; Swaminathan & Schellenberg, 2018; Thomas et al., 2016). In short, online testing did not compromise the reliability and validity of the MET.

Positive associations with scores on the Gold-MSI provide strong evidence of construct validity for the online version of the MET. Previous studies conducted in-person documented that as the degree of musicianship and amount of practice (Wallentin et al., 2010) or duration of music training (Swaminathan & Schellenbergs, 2021) increases, so does performance on the MET. In the present investigation, associations with Music Training as measured by the Gold-MSI were higher than those of Swaminathan et al., (2021) comparison sample. We attribute it to the relatively high variability in music training, and the high proportion of professional musicians tested on-line. Also, we found positive associations between MET scores and other aspects of self-reported musical expertise measured by the Gold-MSI (i.e., Active Engagement, Emotions, Perceptual Abilities, and Singing Abilities). In the Gold-MSI validation study, Müllensiefen et al. (2014) reported a comparable pattern of associations using short beat alignment and melodic memory tasks; our results extended these associations, indicating that musical skills and experience not limited to music lessons or playing an instrument, but instead they are multifaceted. Moreover, even though the musical skills tested by the MET are based on auditory short-term/working memory and perceptual discrimination, performance was predictive of a broad range of musical behaviors and expertise.

We found no association between musical abilities and age of onset of music lessons after duration of music training was held constant, as in the comparison sample. This finding raises the possibility that proposals of plasticity effects arising from early music training (Penhune, 2019, 2020; Penhune & De Villiers-Sidani, 2014) might be exaggerated. Longitudinal evidence in childhood shows that musical ability is independent of music training when levels of musical ability measured

5 years previously are taken into account (Kragness et al., 2021). However, other findings reveal behavioral advantages and structural brain differences as a consequence of early training, even after accounting for duration of training (Bailey et al., 2014; Bailey & Penhune, 2010, 2012, 2013). It could be that early onset of music training explains some musical abilities (such as rhythm synchronization and production abilities), but not other abilities (such as those measured by the MET).

An important advantage of online recruitment is that it allowed reaching a large sample of motivated people, including working musicians or musician-academics. Our sample was also heterogeneous in terms of age and education, which tend to vary minimally when participants are recruited from undergraduate courses in introductory Psychology, as in the MET comparison sample (Swaminathan & Schellenberg, 2021). The status of age and its relation to cognition is more ambiguous, because some abilities, such as processing speed, start to decline relatively early in life, whereas others continue to peak until after age 40 (Hartshorne & Germine, 2015). In any event, age, education and our online measure of cognitive ability were predictive of performance on the MET. In the comparison sample, MET scores correlated positively with three different measures of cognitive ability (digit span forward, digit span backward, and Raven's tests). Thus, as with virtually any specific cognitive ability, individual differences in musical ability vary positively with general ability (Carroll, 1993), whether they are measured in person or online.

Although the association between MET scores and cognitive abilities was consistent with previous research (e.g., Swaminathan et al., 2017, 2018, 2021; Swaminathan & Schellenberg, 2018), and strong even when other variables were held constant (Table 17), cognitive ability was a better predictor of scores on the Rhythm compared to the Melody subtest. In Swaminathan et al. (2021), working memory (as measured by digit span backward) emerged as a better predictor of Rhythm than of Melody scores. On the other hand, music training was a better predictor of Melody subtest with respect to Rhythm subtest in both the online and in-person samples. This difference extended to other aspects of musical expertise measured by the Gold-MSI (i.e., Active Engagement, Perceptual Abilities, Singing Ability, and the General Factor). Therefore, performance on the Melody subtest appears to rely more on individual differences in exposure to music; performance on the Rhythm subtest, instead, seems to be more strongly associated with nonmusical individual differences. As suggested by Swaminathan & Schellenberg (2021), this result might be explained considering that the Rhythm subtest refers to a universal feature of music, while performance on the Melody subtest seems to be more strongly influenced by exposure to pitch structures that are specific to Western music.

Additionally, performance only in the Melody subtest -and not the Rhythm subtest- was linked to a lower level of mind-wandering; however, this association disappeared when other predictors of Melody scores were held constant. In Wang et al., (2015) study, highly trained musicians showed an enhanced ability to sustain attention during a temporal (but not in a visual) discrimination task, remaining evident when cognitive ability was held constant.

The multiple regression analyses we run served to identify which subscales made independent contributions to predicting performance on the MET, considering that the Gold-MSI subscales had considerable overlap. In addition to the Music Training subscale, the Perceptual Abilities subscale was a robust predictor of Melody, Rhythm, and Total scores, and, in the case of Rhythm, even superior to Music Training. These findings highlight participants' meta-cognitive awareness of their musical ability: individual differences in self-reports of music perception skills, measured before taking the MET, correlated with musical abilities measured subsequently and objectively.

In conclusion, our findings revealed that online administration of MET is a valid and reliable alternative to traditional in-person testing of musical abilities.

5. General conclusions

5.1 Summary of the findings

The aim of the present dissertation was to explore further the relationship between music and cognitive abilities over the adult life span, with particular attention to the older adult population, investigating through three studies: a) the effect of listening to music on older adults' affective state and cognitive tasks; b) the relationship between music training and cognitive ability in the adult life span; c) the measurement of musical abilities using an online format.

Study 1 focused on examining whether listening to brief excerpts of music had positive effects on older adults' cognitive abilities and affective state. In the literature, the effect of listening to music in cognition showed mixed results. Although a positive effect of music, for instance, on source memory (e.g., Palumbo et al., 2018), on memory encoding (Ferreri et al., 2014), on episodic memory and processing speed (Bottiroli et al., 2014), was found, on the other, a negative (Giannouli et al., 2019; Hirokawa et al. 2004) or null effect (Borella et al., 2014; Borella et al., 2017) on working memory, and on executive functioning (Giannouli et al., 2019) emerged. Indeed, the reason for such mixed findings might be related to the large individual differences characterizing the older adult population, together with possible contextual factors (e.g., task difficulty). Moreover, executive function deficits in older people (see Braver & West, 2005; Hasher & Zacks, 1988), and their difficulties in inhibiting irrelevant information during encoding and retrieval (Borella et al., 2007), must be considered in order to interpret these heterogeneous results. Effects of music listening on cognition and emotion have been studied when music is played either *while* participants complete a task (i.e., background music) or *before* they begin the task. Considering that background music seems to have a stronger negative effect on older adults' performance, in Study 1 we decided to propose the listening condition following the second presentation modality (i.e., before completing the tasks). Our study was conceptualized in the framework of the arousal and mood hypothesis (Thompson et al., 2001), which is based on the idea that music influences arousal and/or mood, which in turn affects cognitive performance. To our knowledge, no studies have directly tested the arousal and mood hypothesis in a sample of older adults, so we wanted to investigate whether this hypothesis was confirmed also in older people. To do so, we proposed two different musical stimuli, selected for being clearly happy- or sad-sounding, which have been previously used (e.g., Borella et al., 2014, 2019; He et al., 2017; Thompson et al., 2001). Because the literature on the way music listening affects cognitive performance has rarely considered the duration of the exposure to the listening condition as a relevant aspect—rather, the length of the listening condition varied from a few seconds (e.g., Silva et al., 2020) to 10 mins (e.g., Hirokawa, 2004)—we decided to propose two different durations. As for the cognitive tasks, our dependent cognitive measures were tests of executive functions (verbal fluency as well as flexibility and speed), WM, and an arithmetic ability.

The results of this study showed that, as expected, compared to listening to sad-sounding music or to the control condition (a spoken-word recording), listening to happy-sounding music increased arousal levels and improved mood, whereas listening to sad-sounding music was accompanied by reduced levels of arousal. Participants who heard happy-sounding music exhibited improvements in visuospatial WM, in line and consistently with the arousal and mood hypothesis that we wanted to test. However, at the individual level, improvements in arousal and/or mood had no association with improvements in WM. Such a pattern of findings highlighted that in some cases, the individual differences among older adults (Larcom & Isaacowitz, 2009) make correlations between scores on one task and scores on another relatively difficult to observe, in line with previous mood-induction research. Emotional responses to music, in particular, showed that older adults seem to have stronger emotional reactions to happy-sounding music compared to music that conveys other emotions. Moreover, considering older adults' deficits in recognizing positive and negative emotions, the deficits appear to be larger for music that conveys negative emotions, as well as for negative facial expressions (e.g., Sutcliffe et al., 2017). Our results seem to be in line with the "positivity effect" (Carstensen & DeLiema, 2018), showing that, in general, older individuals preferably attend to and remember positive compared to negative stimuli. In addition, these factors might explain why older adults did not show a decrease in mood after listening to the sad-sounding music, even if they showed lower arousal levels. Moreover, the lack of an association between emotional responding and cognitive ability observed here could also be related to differences in measurement approaches.

Regarding the duration of music stimuli, Study 1 did not show significant results, but our analyses documented positive evidence of improvements in WM and in mood for listeners in the Mozart group, with strong evidence for those who heard the *longer* duration. Indeed, because older people generally need more time to process auditory stimuli, and perhaps to feel the effect, they are likely to benefit from longer exposure to the music. This speculation needs to be tested in the future.

To conclude and sum up, older adults who heard the happy-sounding music exhibited larger positive changes in arousal, mood, and WM, as predicted by the arousal and mood hypothesis (Thompson et al., 2001). Nevertheless, the independence between measures of emotion and cognition appears to stem from large inter- and intraindividual variability that characterizes the whole aging process.

An overview of the content and main results of Study 1 is presented in Table 20.

Table 20. *Overview of the content and main results of Study 1.*

General aim(s)	Sample	Outcomes	Main results
Study 1: Exploring the effect of listening to music on older adults' affective state and cognitive tasks	132 Healthy older adults (age range: 65-75 years)	*Affective state*: Mood and arousal: SAM (Bradley et al., 1994) *Cognitive tasks:* - VS-WM: backward Corsi Blocks Task (Corsi, 1972) - EF: Verbal Fluency (Novelli et al., 1986), TMT-B (Amodio et al., 2002) - Arithmetical abilities: AC-FL (Caviola et al., 2016).	WM: + in happy-sounding music condition vs. sad-sounding and control group Arousal and mood: + in happy-sounding condition + all cognitive tasks from pre- to post-test No significant interaction for duration, but + WM and mood for happy-sounding longer condition

Study 2 addressed the issue of improving knowledge about the link between music training and cognitive ability across the adult life span. A growing body of research has investigated this relationship, seeing that music training has positive associations with tasks related to music, visuospatial, language, and general cognitive abilities tasks' (see Swaminathan & Schellenberg, 2019). Indeed, evidence seems to suggest that musically trained participants outperform their untrained counterparts on a variety of music cognition tasks, and—coherently with constant practice—musically trained participants seem to be very good listeners (Kraus & Chandrasekaran, 2010; Strait et al., 2011). However, as concerns cognitive abilities, considering that most of the studies in this field are quasi-experimental or correlational, precluding inferences of causation, it is reasonable to assume that performing music might improve a person's listening ability, but the reverse causal direction is equally plausible (Schellenberg et al., 2013, 2019). Moreover, high-functioning individuals usually begin activities—music training included—earlier, and they are more likely to take music lessons for many years. Hence, stating that musicians are generally "smarter" sounds a bit speculative.

An issue that makes literature on this topic confusing concerns the criteria for defining a professional musician, which often follow the "six-year rule", that is, having at least 6 years of musical expertise (Zhang et al., 2020). Indeed, it is worth mentioning that this rule ignores whether individuals are working or have worked as musicians, failing in distinguishing real professionals (i.e., persons whose daily behaviors are or were dedicated to music) from musically trained individuals who become workers in other disciplines. Evidence also shows that personality traits predict occupational choices (Holland, 1997). In particular, openness to experience seems to be associated with creativity, and to predict musical behaviors and skills (Corrigal et al, 2013; Lima et al., 2020).

In Study 2, we examined group differences in musical ability, personality and general cognitive ability in a sample of professional musicians (who had a music-related job and/or were enrolled as students in a university-level music program), in participants who were musically trained (at least 6 years of music lessons but did not meet the criteria for professionals) and untrained participants (who had a maximum of 2 years of music training).

In line with our expectations and previous evidence, the finding of Study 2 was that on personality measures, professional musicians and musically trained participants scored similarly, but they scored higher than untrained participants did on agreeableness, openness-to-experience, and the personality stability meta-trait. The professional musicians also scored higher than the other two groups did on extraversion and the engagement meta-trait. Surprisingly, on cognitive ability, professionals were indistinguishable from untrained participants, while musically trained participants—not professionals—exhibited the highest cognitive ability. This result seems to suggest that—differently from how they are generally considered—professional musicians were not "smarter", and they did not show better cognitive performance with respect to trained participants and untrained people.

As for Study 3, considering that internet is now part of our daily life, and because the development of a number of online platforms provided the possibility to recruit and test a considerable number of people (e.g., Gosling & Mason, 2015; Grootswagers, 2020), we aimed to determine whether an objective test of musical ability (i.e., the MET), could be successfully administered online.

Indeed, online studies have some features that make them even better than in-person testing (e.g., Casler et al., 2013; Dandurand et al., 2008; Gosling et al., 2004), even though there are some concerns about online format, such as lack of control over characteristics of the samples and testing contexts (e.g., Birnbaum, 2004; Krantz & Dalal, 2000). In particular, internet samples can be more representative of the general population, and recruiting people as well as collecting data can be more efficient in terms of time and cost. Moreover, participants may feel more relaxed at their home than they would in a laboratory. Additionally, accessing a specific target, such as musicians, is simpler. Despite these advantages, online testing needs cautious experimental designs in order to maximize control (e.g., Gosling et al., 2004) and to engage people, so that they complete the whole task or a whole experiment. In addition, the contexts of online testing less controlled in terms of extraneous sounds, technical aspects of stimulus presentation, and potential interruptions (e.g., Milne et al., 2020)—although this variability can be reduced by asking participants to follow specific instructions (e.g., to wear headphones)—so auditory and music research can be particularly challenging.

In Study 3, participants were tested with an online version of the MET, which includes a Melody and a Rhythm subtest, requiring participants to determine whether two auditory sequences were or not identical. In addition, we proposed the Goldsmiths Musical Sophistication Index (Gold-MSI), a test of general cognitive ability, and self-report questionnaires that measured basic demographics, mind-wandering, and personality.

The study found that results from the online format were similar to those from in-person testing. In particular, the internal reliability of the MET was maintained, and strong associations with the Gold-MSI confirmed construct validity. Moreover, our results were in line with previous literature concerning correlations with other measures (openness to experience, cognitive ability, and mind-wandering), as predicted. Therefore, online administration of the MET proved to be a reliable and valid way to measure musical ability.

An overview of the content and main results of Study 2 and 3 is presented in Table 21.

Table 21. *Overview of the content and main results of Study 2 and 3.*

General aim(s)	Sample	Outcomes	Main results
Study 2: Investigate the relationship between music training and cognitive ability in the adult life span	608 adults (age range: 18-88 years)	*Objective behavioral tests* Musical ability: Musical Ear Test (MET; Wallentin et al., 2010) General cognitive ability: Matrix Reasoning Item Bank (MaRs-IB; Chierchia et al., 2019) *Questionnaires:* Musical expertise: Gold-MSI (Müllensiefen et al., 2014) Personality: Big Five Inventory (BFI, John & Srivastava, 1999) Mind-Wandering Questionnaire (MWQ, Mrazek et al., 2013)	Professional musicians: > objective (MET) and self-report measures (Gold-MSI) of musical ability. Professional musicians & musically trained participants: > untrained participants on agreeableness, openness-to-experience, and stability metatrait. Professional musicians > musically trained, untrained participants on extraversion and engagement metatrait. Professional musicians = untrained participants in cognitive ability Musically trained nonprofessionals: > highest cognitive ability

Study 3: Determine whether an objective test of musical ability could be successfully administered online	642 adults (age range: 18-84 years)	*Objective behavioral tests* Musical ability: Musical Ear Test (MET; Wallentin et al., 2010) General cognitive ability: Matrix Reasoning Item Bank (MaRs-IB; Chierchia et al., 2019) *Questionnaires:* Musical expertise: Gold-MSI (Müllensiefen et al., 2014) Personality: Big Five Inventory (BFI, John & Srivastava, 1999) Mind-Wandering Questionnaire (MWQ, Mrazek et al., 2013)	Online findings similar to in-person testing: (1) internal reliability of the MET maintained (2) construct validity: confirmed by strong associations with Gold-MSI scores (3) correlations with other measures (e.g., openness to experience, cognitive ability, mind-wandering) as predicted (4) mean levels of performance similar for individuals with no music training, (5) musical sophistication: better predictor of performance on the Melody than on the Rhythm subtest.

Overall, results from Study 1 show that listening to happy-sounding music can improve older adults' mood and WM performance. These findings have important implications in relation to age-related and individual differences in listening to music and cognitive abilities. In fact, results from Study 1 point to the importance of considering listening to music to improve older adults' affective state as a means to influence cognitive performance. In addition, our results, which are partially in line in confirming the "arousal and mood hypothesis", contribute to the growing literature on the effect of listening to music on older adults' affective states and cognitive abilities. For older people, music listening may be a means to induce positive affective states, and to improve cognitive ability.

The pattern of findings from Study 2 supports previous suggestions regarding the role of personality in professional musicians, documenting important differences between professional musicians and nonprofessional but musically trained individuals. Moreover, the findings have implications in studying the link between music training and cognition, because professionals did not show better cognitive performance with respect to musically trained participants and participants without music training. However, these differences need to be considered carefully when interpreting the results of published research, and when designing future studies.

As for Study 3, this last has revealed that online administration of MET is a valid and reliable alternative to traditional in-person testing of musical abilities. Strong associations between the accuracy on the MET and musical sophistication and training, especially for the Melody subtest, were also consistent with studies using in-person testing of MET (Swaminathan et al., 2021). Considering the expansion of access to the internet, this study contributes to the growing literature on the utility of online measurement as an alternative, or complement, to laboratory testing for psychological research.

5. 2 Limitations and recommendations for future directions

Despite the interesting findings emerging from these studies, some issues emerge when we consider their limitations. Regarding Study 1, possible limitations of the study include the limited sample size and the between-subjects design, which did not allow us to investigate the effect of listening to the different excerpts on the same participants. Moreover, our exclusive use of self-report measures could also be complemented with psychophysiological measures (e.g., skin conductance), which would be useful to examine further the effect of music listening among older adults.

As for Study 2, a possible limitation refers to the cognitive ability measure we chose: in fact, even though matrix-reasoning tests are considered to be the best single-test proxy for *g* factor (e.g., Deary & Smith, 2004), the results could be test-specific (or specific to tests of matrix reasoning). Our choice of test was motivated by the testing context and the online format, but future studies are needed to confirm these findings.

Concerning Studies 2 and 3, which were performed online, it must be said that online testing occurs in contexts that are more variable and less controlled in terms of extraneous sounds, technical aspects of stimulus presentation, and potential interruptions (e.g., Milne et al., 2020). Therefore, even though this variability can be reduced by asking participants to follow specific instructions (e.g., to wear headphones), experimental control remains limited. In fact, even if participants are specifically asked to perform the experiment in a quiet environment and to avoid distractions, internet testing makes it difficult to control for extraneous sounds, which represents a challenge for testing in general, and for auditory research in particular.

In accordance with these considerations, future studies would benefit from taking into account these aspects. Moreover, considering that studies that include older people are still insufficient, future research should involve older participants, to shed additional light regarding the relationship between music and cognition in older age. In addition, some of the issues addressed in this dissertation are also worth investigating in relation to pathological aging, considering the power of music to influence the emotional state, and cognitive abilities.

Finally, future studies might also focus also on older adults' cognitive reserve, that is, a resource to be used when the brain experiences increased burden (Stern et al., 2003). In this framework, musical training could represent an important factor that could protect against cognitive decline.

5.3 Conclusions

In conclusion, the present dissertation set out the relationship between music and cognition.

In particular, the benefits of listening to music on affective state and on cognition in healthy older adults were further explored. The study showed that listening to a happy-sounding music had positive effects on older adults' arousal, mood, and WM. Some cognitive domains (i.e., executive functions, arithmetic ability), however, seemed to be less affected by music listening with respect to others (i.e., WM). Therefore, these results highlight the large inter- and intraindividual variability characterizing the aging process (Fagot et al., 2018), which requires further research.

Moreover, this dissertation advances our understanding regarding professional musicians' "profile" and proposes specific criteria to define what a professional musician is, considering not only the years of study, but also the person's main job. Our findings showed that professional musicians reported higher scores in openness to experience, and extraversion traits across the adult life span. Importantly, professional musicians did not show better cognitive abilities with respect to untrained persons and musically trained nonprofessionals -who reported the highest scores. Nevertheless, professionals reported the highest scores on objective musical ability measures, highlighting their proficiency in auditory/musical tasks. Hence, professional musicians showed better listening abilities, but they did not report better cognitive abilities in general, so the widespread idea to consider this particular population "smarter" needs to be taken with caution.

Finally, this work contributes to elucidating the possibility of using online methods to test musical abilities, which emerged as a reliable and valid way to reach a large number of individuals, from an ample range of ages and a specific population (i.e., musicians).

Although further research is needed, such findings would guide future studies aiming to investigate comprehensively the relationship between music and cognition, as well as to explore the use of music listening to sustain older adults’ affective states, emotions, and cognitive abilities. Additionally, this dissertation represents a prompt in better exploring the benefits of music training on cognitive abilities across the adult life span and in aging. Finally, it highlights the importance of using new technology as a means to obtain information concerning populations that would be difficult to reach without using the internet.

References

Amer, T., Kalender, B., Hasher, L., Trehub, S. E., & Wong, Y. (2013). Do older professional musicians have cognitive advantages?. *PloS one*, *8*(8), e71630. https://doi.org/10.1371/journal.pone.0071630

Amodio, P., Wenin, H., Del Piccolo, F., Mapelli, D., Montagnese, S., Pellegrini, A., Musto, C., Gatta, A. & Umiltà, C. (2002). Variability of trail making test, symbol digit test and line trait test in normal people. A normative study taking into account age-dependent decline and sociobiological variables. *Aging Clinical and Experimental Research*, *14*(2), 117-131. https://doi.org/10.1007/BF03324425

Andrews, M. W., Dowling, W. J., Bartlett, J. C., & Halpern, A. R. (1998). Identification of speeded and slowed familiar melodies by younger, middle-aged, and older musicians and nonmusicians. *Psychology and Aging*, *13*, 462-471. https://doi.org/10.1037/0882-7974.13.3.462

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Arnaud, L., Lemaire, P., Allen, P., Michel, B. F. (2008). Strategic aspects of young, healthy older adults', and Alzheimer patients' arithmetic performance. *Cortex*, *44*(2), 119-130. https://doi.org/10.1016/j.cortex.2006.03.001

Asztalos, K., & Csapó, B. (2014). Online assessment of musical abilities in Hungarian primary schools—results of first, third and fifth grade students. *Bulletin of the International Kodály Society*, *39*(1), 3–14.

Bailey, J. A., & Penhune, V. B. (2010). Rhythm synchronization performance and auditory working memory in early- and late-trained musicians. *Experimental Brain Research*, *204*, 91–101. https://doi.org/ 10.1007/s00221-010-2299-y

Bailey, J. A., & Penhune, V. B. (2012). A sensitive period for musical training: Contributions of age of onset and cognitive abilities. *Annals of the New York Academy of Sciences*, *1252*, 163–170. https://doi.org/10.1111/j.1749-6632.2011.06434.x

Bailey, J. A., & Penhune, V. B. (2013). The relationship between the age of onset of musical training and rhythm synchronization performance: Validation of sensitive period effects. *Frontiers in Auditory Cognitive Neuroscience*, *7*, 227. https://doi.org/10.3389/fnins.2013.00227

Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2014). Early musical training: Effects on auditory motor integration and grey matter structure in ventral premotor cortex. *Journal of Cognitive Neuroscience*, *26*(4), 755–767. https://doi.org/10.1162/jocn_a_00527

Balbag, M. A., Pedersen, N. L., & Gatz, M. (2014). Playing a musical instrument as a protective factor against dementia and cognitive impairment: A population-based twin study. *International Journal of Alzheimer's Disease*, *2014*(836748). https://doi.org/10.1155/2014/836748

Bentley, A. (1966). *Measures of Musical Abilities*. London: Harrap.

Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*, *25*, 399-410.

Birnbaum, M. H. (2004). Methodological and ethical issues in conducting social psychological research via the Internet. In C. Sansone, C. C. Morf, & A. T. Panter (Eds.), *Handbook of Methods in Social Psychology* (pp. 359–382). Sage Publications, Inc.

Borella E., Carretti B., Meneghetti C., Carbone E., Vincenzi M., Madonna J.C., Grassi M., Fairfield B., & Mammarella N. (2019). Is working memory training in older adults sensitive to music? *Psychological Research*, *83*(6):1107-1123. https://doi.org/10.1007/s00426-017-0961-8.

Borella, E., Carretti, B., & Mammarella, I. (2006). Do working memory and susceptibility to interference predict individual differences in fluid intelligence? *European Journal of Cognitive Psychology*, *18*(1), 51-69. https://doi.org/10.1080/09541440500215962

Borella, E., Carretti, B., & De Beni, R. (2008). Working memory and inhibition across the adult life-span. *Acta Psychologica*, *128*(1), 33-44. https://doi.org/10.1016/j.actpsy.2007.09.008

Borella, E., Carretti, B., Grassi, M., Nucci, M., & Sciore, R. (2014). Are age-related differences between young and older adults in an affective working memory test sensitive to the music effects?. *Frontiers of Aging Neuroscience*, *6*, 298. https://doi.org/10.3389/fnagi.2014.00298

Bottiroli, S., Rosi, A., Russo, R., Vecchi, T., & Cavallini, E. (2014). The cognitive effects of listening to background music on older adults: processing speed improves with upbeat music, while memory seems to benefit from both upbeat and downbeat music. *Frontiers of Aging Neuroscience*, *6*, 284. https://doi.org/10.3389/fnagi.2014.00284

Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: The self- assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49-59. https://doi.org/10.1016/0005-7916(94)90063-9

Braver, T. S., & West, R. (2008). Working memory, executive control, and aging. In Craik F. I. M.; Salthouse, T. A. (Eds.), *The hand-book of aging and cognition* (3rd ed.) (pp. 311–372), Psychology Press. https://doi.org/10.1080/01924780903295796

Buckner, R. L. (2004). Memory and executive function in aging and AD: multiple factors that cause decline and reserve factors that compensate. *Neuron*, *44*(1), 195-208. https://doi.org/10.1016/j.neuron.2004.09.006

Butkovic, A., & Modrusan, I. (2021). Personality differences among musicians: Real differences or stereotypes?. *Psychology of Music*, *49*(2), 216-226.

Butkovic, A., Ullén, F., & Mosing, M. A. (2015). Personality related traits as predictors of music practice: Underlying environmental and genetic influences. *Personality and Individual Differences*, *74*, 133-138.

Buttsworth, L. M., & Smith, G. A. (1995). Personality of Australian performing musicians by gender and by instrument. *Personality and Individual Differences*, *18*(5), 595-603.

Cacciafesta, M., Ettorre, E., Amici, A., Cicconetti, P., Martinelli, V., Linguanti, A. E. E. A., ... & Marigliano, V. (2010). New frontiers of cognitive rehabilitation in geriatric age: the Mozart Effect (ME). *Archives of gerontology and geriatrics*, *51*(3), e79-e82.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404–431. https://doi .org/10.1037/0033-295X.97.3.404

Carretti, B., Cornoldi, C., Caldarola, N., & Tencati, C. (2013). CO-TT, Comprensione Orale Test e Trattamento [Oral Comprehension, Assessment and Training]. Erickson: Trento, Italy.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press. https://doi.org/10.1017/CBO978051 1571312

Carstensen, L. L., & DeLiema, M. (2018). The positivity effect: A negativity bias in youth fades with age. *Current Opinion in Behavioral Sciences*, *19*, 7–12. https://doi.org/10.1016/j.cobeha.2017.07.009

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Psychology*, *29*(6), 2156–2160. https://doi.org/10.1016/j.chb.2013. 05.009

Caviola, S., Gerotto, G., Lucangeli, D., & Mammarella, I. C. (2016). Test AC-FL: Prove di fluenza per le abilità di calcolo [*AC-FL Test: Math Fluency abilities test*]. Erickson: Trento, Italy.

Chabris, C.F. (1999). Prelude or requiem for the "Mozart Effect"? *Nature*, *400*, 826. https://doi.org/10.1038/23608

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in cognitive sciences*, *17*(4), 179-193. https://doi.org/10.1016/j.tics.2013.02.007

Chetverikov, A., & Upravitelev, P. (2015). Online versus offline: The Web as a medium for response time data collection. *Behavior Research Methods*, *48*(3), 1086–1099. https://doi.org/10.3758/ s13428-015-0632-x

Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S. J. (2019). The matrix reasoning item bank (MaRs-IB): Novel, open-access abstract reasoning items for adolescents and adults. *Royal Society Open Science*, *6*(10), 190232. https://doi.org/ 10.1098/rsos.190232

Chow, R., Noly-Gandon, A., Moussard, A., Ryan, J. D., & Alain, C. (2021). Effects of transcranial direct current stimulation combined with listening to preferred music on memory in older adults. *Scientific Reports*, *11*(1), 1-13. https://doi.org/10.1038/s41598-021-91977-8

Connelly, S. L., Hasher, L., & Zacks, R. T. (1991). Age and reading: the impact of distraction. *Psychology and aging, 6*(4)*,* 533. https://doi.org/10.1037/0882-7974.6.4.533

Cooper, P. K. (2019). It's all in your head: A meta-analysis on the effects of music training on cognitive measures in schoolchildren. *International Journal of Music Education*, *38*(3), 321–336. https:// doi.org/10.1177/0255761419881495

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in Psychology*, *4*, 00222. https://doi.org/10.3389/fpsyg.2013.00222

Corrigall, K. A., & Schellenberg, E. G. (2015). Predicting who takes music lessons: Parent and child characteristics. *Frontiers in Psychology*, *6*(282). https://doi.org/10.3389/fpsyg.2015. 00282

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. Ph.D. Unpublished Thesis, McGill University, Montreal, Canada.

Craik, F. & Salthouse, T. (2008). *The Handbook of Aging and Cognition*. 3rd Edn. Psychology Press. https://doi.org/10.1002/acp.1505

Cribb, C., & Gregory, A. H. (1999). Stereotypes and personalities of musicians. *The Journal of Psychology*, *133*(1), 104-114. https://doi.org/10.1080/00223989909599725

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*(3), B1-B10. https://doi.org/10.1016/s0010-0277(00)00136-0

Dandurand, F., Shultz, T. R., & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods*, *40*(2), 428–434. https://doi.org/10.3758/brm. 40.2.428

De Beni, R., Borella, E., Carretti, B., Marigo, C., & Nava, L. A. (2008). *BAC. Portfolio per la valutazione del benessere e delle abilità cognitive nell'età adulta e avanzata [The Assessment of Well-being and Cognitive Abilities in Adulthood and Aging]*. Giunti OS: Firenze, Italy.

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners. *Psychological Science*, *27*(6), 848–858. https://doi.org/10.1177/0956797616639301

Degé, F. (2021). Music lessons and cognitive abilities in children: How far transfer could be possible. *Frontiers in Psychology*, *11*, 557807. https:// doi.org/10.3389/fpsyg.2020.557807

Degé, F., Kubicek, C., & Schwarzer, G. (2011). Music lessons and intelligence: A relation mediated by executive functions. *Music Perception*, *29*(2), 195–201. https://doi.org/10.1525/mp.2011.29.2.195

Deliège, I., & Sloboda, J. A. (1997). *Perception and cognition of music*. Deliège, I., & Sloboda, J. A. (Eds.). Psychology Press. https://doi.org/10.4324/9780203344262

DeYoung, C. G. (2013). The neuromodulator of exploration: A unifying theory of the role of dopamine in personality. *Frontiers in Human Neuroscience*, *7*, 762. https://doi.org/10.3389/fnhum.2013.00762

Dickson, G. T., & Schubert, E. (2019). How does music aid sleep? Literature review. *Sleep medicine*, *63*, 142-150. https://doi.org/10.1016/j.sleep.2019.05.016

Drake, R. M. (1954). *Drake Musical Aptitude Tests*. Science Research Associates.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363-406.

Fagot, D., Mella, N., Borella, E., Ghisletta, P., Lecerf, T., & De Ribaupierre, A. (2018). Intra-individual variability from a lifespan perspective: A comparison of latency and accuracy measures. *Journal of Intelligence*, *6*(1), 16. https://doi.org/10.3390/jintelligence6010016

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, *2*(4), 290–309. https://doi.org/10.1207/s15327957pspr0204_5

Feist, G. J. (2019). Creativity and the Big Two model of personality: Plasticity and stability. *Current Opinion in Behavioral Sciences*, *27*, 31–35. https://doi.org/10.1016/j.cobeha.2018.07.005

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports*, *11*, 1382. https://doi.org/10.1038/s41598-020-80866-1

Ferreri, L.; Bigand, E.; Perrey, S.; Muthalib, M.; Bard, P.; Bugaiska, A. (2014). Less effort, better results: how does music act on prefrontal cortex in older adults during verbal encoding? An fNIRS study. *Frontiers of Human Neuroscience*, *8*, 301. https://doi.org/10.3389/fnhum.2014.00301

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: a unified JavaScript framework for online studies. In *International Conference on Computational Social Science* IC2S, July 2016.

Flaten, M. A., Asli, O., & Simonsen, T. (2006). The effect of stress on absorption of acetominophen. *Psychopharmacology*, *185*, 471-478. https://doi.org/10.1007/s00213-006-0324-4

Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*, *47*(1), 72-89. https://doi.org/10.1038/s41386-021-01132-0

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., (2009). Universal recognition of three basic emotions in music. *Current Biology*, *19*, 573-576. doi:10.1016/j.cub.2009.02058

Fujii, S., & Schlaug, G. (2013) The Harvard Beat Assessment Test (H- BAT): A battery for assessing beat perception and production and their dissociation. *Frontiers in Human Neuroscience*, *7*, 771. https:// doi.org/10.3389/fnhum.2013.00771

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 223–248). Oxford University Press.

Gallina, P., Saugo, M., Antoniazzi, M., Fortuna, P., Toffanin, R., Maggi, S., & Benetollo, P. P. (2006). Validazione della scheda per la valutazione multidimensionale dell'anziano (SVAMA) (Validation of a multidimensional assessment profile for older adults). *Tendenze Nuove*, *3*, 229–263. 19. https://doi.org/10.1450/22126

Gastón, E. T. (1957). *Test of Musicality*. Odell's Instrumental Service.

George, J. M., & Zhou, J. (2001). When openness to experience and conscientiousness are related to creative behavior: an interactional approach. *Journal of applied psychology*, *86*(3), 513-524. https://doi.org/10.1037/0021-9010.86.3.513

Giannouli, V., Kolev, V., & Yordanova, J. (2019). Is there a specific Vivaldi effect on verbal memory functions? Evidence from listening to music in younger and older adults. *Psychology of Music*, *47*(3), 325-341. https://doi.org/10.1177/0305735618757901

Gjermunds, N., Brechan, I., Johnsen, S., & Watten, R. G. (2020). Personality traits in musicians. *Current Issues in Personality Psychology*, *8*(2), 100-107. https://doi.org/10.5114/cipp.2020.97314

Goldberg L.R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26-34. https://doi.org/10.1037/0003-066X.48.1.26

Gordon, E. (1965). Musical aptitude profile: Manual. Houghton Mifflin.

Gordon, E. (1979). Developmental music aptitude as measured by the Primary Measures of Music Audiation. *Psychology of Music*, *7*(1), 42-49. https://doi.org/10.1177/030573567971005

Gordon, E. (1989). Advanced in Measures of Music Audiation. GIA Publications.

Gosling, S. D., & Mason, W. (2015). Internet research in psychology. *Annual Review of Psychology*, *66*(1), 877–902. https://doi.org/10.1146/annurev-psych-010814-015321

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, *59*(2), 93–104. https://doi.org/10.1037/0003-066x.59.2.93

Grassi, M., Meneghetti, C., Toffalini, E., & Borella, E. (2017). Auditory and cognitive performance in elderly musicians and nonmusicians. *PLoS One*, *12*(11), e0187881.

Greenberg, D. M., Müllensiefen, D., Lamb, M. E., & Rentfrow, P. J. (2015). Personality predicts musical sophistication. *Journal of Research in Personality*, *58*, 154–158. https://doi.org/10.1016/j. jrp.2015.06.002

Groarke, J. M. & Hogan, M. J. (2018). Development and psychometric evaluation of the adaptive functions of music listening scale. *Frontiers in Psychology*, *9*, 516. https://doi.org/10.3389/fpsyg.2018.00516

Groarke, J. M. & Hogan, M. J. (2019). Listening to self-chosen music regulates induced negative affect for both younger and older adults. *PLoS One*, *14*(6), e0218017. https://doi.org/10.1371/journal.pone.0218017

Grootswagers, T. (2020). A primer on running human behavioural experiments online. *Behavior Research Methods*, *52*(6), 2283–2286. https://doi.org/10.3758/s13428-020-01395-3

Hallam, S. (2009). The Effects of Listening to Music on Children's Spatial Task Performance. *British Psychological Society Education Review*, *25*(2), 22-26. https://doi.org/10.1177/0305735607068885

Halpern, A. R., Kwak, S., Bartlett, J. C., & Dowling, W. J. (1996). Effects of aging and musical experience on the representation of tonal hierarchies. *Psychology and Aging*, *11*, 235-246. https://doi.org/10.1037/0882-7974.11.2.235

Hanna-Pladdy, B., & Gajewski, B. (2012). Recent and past musical activity predicts cognitive aging variability: direct comparison with general lifestyle activities. *Frontiers in Human Neuroscience*, *6*, 198. https://doi.org/10.3389/fnhum.2012.00198

Hanna-Pladdy, B., & MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology*, *25*, 378-386. https://doi.org/10.1037/a0021895

Hasher, L., & Zacks, R. (1988). Working memory, comprehension, and aging: A review and a new view. In Bower, G., Ed., *The psychology of learning and motivation* (pp. 193–225). Academic Press. https://doi.org/10.1016/S0079-7421(08)60041-9

Harris, J. R. (2006). *No two alike: Human nature and human individuality*. Norton.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive function- ing peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, *26*(4), 433–443. https://doi .org/10.1177/0956797614567339

Haynes, B. I., Bauermeister, S., & Bunce, D. A (2017). Systematic review of longitudinal associations between reaction time intraindividual variability and age-related cognitive decline or impairment, dementia, and mortality. *Journal of the International Neuropsychological Society*, *23*(5), 431-445. https://doi.org/10.1017/S1355617717000236

He W.J., Wong W.C., & Hui A.N.N. (2017). Emotional reactions mediate the effect of music listening on creative thinking: Perspective of the arousal-and-mood hypothesis. *Frontiers in Psychology, 8*, 1680. https://doi.org/10.3389/fpsyg.2017.01680

Heilman, K. M., Nadeau, S. E., & Beversdorf, D. O. (2003). Creative innovation: possible brain mechanisms. *Neurocase*, *9*(5), 369-379. https://doi.org/10.1076/neur.9.5.369.16553

Hetland, L. (2000). Learning to make music enhances spatial reasoning. *Journal of aesthetic education*, *34*(3/4), 179-238. https://doi.org/10.2307/3333643

Hirokawa, E. (2004). Effects of music listening and relaxation instructions on arousal changes and the working memory task in older adults. *Journal of Music Therapy, 41*(2), 107-127. https://doi.org/10.1093/jmt/41.2.107.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

Houben, K., & Wiers, R. W. (2008). Measuring implicit alcohol associations via the Internet: Validation of Web-based implicit association tests. *Behavior Research Methods*, *40*(4), 1134–1143. https://doi.org/10.3758/brm.40.4.1134

Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: reality or myth. *Behavioral and Brain Sciences*, *21*(3), 399-407. https://doi.org/10.1017/s0140525x9800123x

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music perception*, *20*(2), 151-171. https://doi.org/10.1525/mp.2002.20.2.151

International Telecommunication Union (2020. Measuring digital development. Facts and figures 2020. Geneva: ITU.

Jäncke, L. (2009). Music drives brain plasticity. *F1000 Biology Reports*, *1*(10), 78. https://doi.org/10.3410/B1-78

Jarosz, A., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *Journal of Problem Solving*, 7(1), 2–9. https://doi.org/10.7771/1932-6246.1167

JASP Team. JASP (Version 0.16.3) 2022 [Computer software].

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Jenkins, J. S. (2001). The Mozart effect. *Journal of the royal society of medicine*, *94*(4), 170-172. https://doi.org/10.1177/014107680109400404

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102–138). Guilford Press.

Juslin, P. N. (2019). Emotional reactions to music: Mechanisms and modularity. In: A. Ferreira Correa (Ed.), *Music, speech, and mind* (pp. 17-51). Curitiba: Brazilian Association of Cognition and Musical Arts

Juslin, P. N., & Sloboda, J. A. (2001). *Music and emotion: Theory and research*. Juslin, P. N., & Sloboda, J. A. (Eds), Oxford University Press.

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: listener, music, and situation. *Emotion*, *8*(5), 668. https://doi.org/10.1037/a0013505

Kämpfe, J., Sedlmeier, P., & Renkewitz, F. (2011). The impact of background music on adult listeners: A meta-analysis. *Psychology of music*, *39*(4), 424-448. https://doi.org/10.1177/0305735610376261

Karma, K. (2007). Musical aptitude definition and measure validation: Ecological validity can endanger the construct validity of musical aptitude tests. *Psychomusicology: A Journal of Research in Music Cognition, 19*(2), 79–90. https://doi.org/10.1037/h0094033

Karwowski, M., & Lebuda, I. (2016). The big five, the huge two, and creative self-beliefs: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(2), 214-232. https://doi.org/10.1037/aca0000035

Kass, R. E., Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773-795. https://doi.org/10.1080/01621459.1995.10476572

Kiss, L. & Linnell, K. J. (2022) Making sense of background music listening habits: An arousal and task-complexity account. *Psychology of Music*, 0(0). https://doi.org/10.1177/03057356221089017

Kragness, H. E., Swaminathan, S., Cirelli, L. K., & Schellenberg, E. G. (2021). Individual differences in musical ability are stable over time in childhood. *Developmental Science*, *24*(4), e13081. https://doi.org/10 .1111/desc.13081

Krantz, J. H., & Dalal, R. (2000). Validity of Web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35–60). Academic Press.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*, 599-605. https://doi.org/10.1038/nrn2882

Kuckelkorn, K. L., de Manzano, O., & Ullén, F. (2021). Musical expertise and personality—Differences related to occupational choice and instrument category. *Personality and Individual Differences*, *173*, 110573. https://doi.org/10.1016/j.paid.2020.110573

Kwalwasser, J., & Dykema, P. W. (1930). K-D MUSIC Tests; Manual of Directions for Victor Records. Nos. 302, 303, 304, 305 and 306, Standard Test Blanks, Set of Matrices (p. 31). Carl Fischer, Inc.

Lang, P. J., Öhman, A., Vaitl, D. (1988). The international affective picture system. The Center for Research in Psychophysiology: University of Florida, US.

Larcom, M. J. & Isaacowitz, D. M. (2009). Rapid emotion regulation after mood induction: Age and individual differences. *The journals of gerontology. Series B, Psychological Sciences*, *64B*(6), 733–741. https://doi.org/10.1093/geronb/gbp077

Laukka, P., & Juslin, P. N. (2007). Similar patterns of age-related differences in emotion recognition from speech and music. *Motivation and Emotion*, *31*(3), 182–191. https://doi.org/10.1007/s11031-007-9063-z

Law, L. N., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PloS one*, *7*(12), e52508. https://doi.org/10.1371/journal.pone.0052508

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. https://doi.org/10.1037/a0024521

Lima, C. F., Correia, A. I., Müllensiefen, D., & Castro, S. L. (2020). Goldsmiths Musical Sophistication Index (Gold-MSI): Portuguese version and associations with socio-demographic factors, personality and music preferences. *Psychology of Music*, *48*(3), 376-388. https://doi.org/10.1177/0305735618801997

Linnemann, A., Wenzel, M., Grammes, J., Kubiak, T. & Nater, U. M. (2018). Music listening and stress in daily life—a matter of timing. *International Journal of Behavioral Medicine*, *25*(2), 223-230. https://doi.org/10.1007/s12529-017-9697-5.

López-González, M., & Limb, C. J. (2012). Musical creativity and the brain. *Cerebrum*, *2012*, 2.

Mammarella, N.; Fairfield, B.; Cornoldi, C. (2007). Does music enhance cognitive performance in healthy older adults? The Vivaldi effect. *Aging Clinical and Experimental Research*, *19*(5), 394-399. https://doi.org/10.1007/BF03324720.

Marimon, M., Hofmann, A., Veríssimo, J., Männel, C., Friederici, A. D., Höhle, B., & Wartenburger, I. (2021). Children's Learning of Non-adjacent Dependencies Using a Web-Based Computer Game Setting. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.734877

Matarazzo, J. D. (1990). Psychological assessment versus psychology testing: Validation from Binet to the school, clinic, and courtroom. *American Psychologist, 45,* 999–1017. https://doi.org/10.1037/0003-066X.45.9.999

McCabe, D. P., Roediger, III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology* , *24*(2), 222-243. https://doi.org/10.1037/a0017619

McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of personality and social psychology*, *52*(6), 1258-1265. https://doi.org/10.1037/0022-3514.52.6.1258

McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. *American psychologist*, *52*(5), 509-516. https://doi.org/10.1037/0003-066X.52.5.509

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McCrae, R. R., Costa, P. T., de Lima, M. P., Simões, A., Ostendorf, F., Angleitner, A., ... & Piedmont, R. L. (1999). Age differences in personality across the adult life span: parallels in five cultures. *Developmental psychology*, *35*(2), 466. https://doi.org/10.1037//0012-1649.35.2.466

Mehrabian, A. & Russell, J. A. (1974). *An approach to environmental psychology*. The MIT Press.

Merriam-Webster Disctionary (2022). “Music”, Retrieved July 8th, 2022, from https://www.merriam-webster.com/dictionary/music

Miles, C. C. (1926). *The early mental traits of three hundred geniuses*. Stanford University Press.

Milne, A. E., Bianco, R., Poole, K. C., Zhao, S., Oxenham, A. J., Billig, A. J., & Chait, M. (2020). An online headphone screening test based on dichotic pitch. *Behavior Research Methods*, *53*(4), 1551-1562. https://doi.org/10.3758/s13428-020-01514-0

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, *22*(11), 1425–1433. https://doi.org/10.1177/0956797611416999

Mrazek, M. D., Phillips, D. T., Franklin, M. S., Broadway, J. M., & Schooler, J. W. (2013). Young and restless: Validation of the Mind- Wandering Questionnaire (MWQ) reveals disruptive impact of mind- wandering for youth. *Frontiers in Psychology*, *4*, 560. https://doi.org/10 .3389/fpsyg.2013.00560

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS ONE*, *9*(2), e89642. https://doi.org/10.1371/journal.pone.0089642

Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., & Velasquez, R. (2004). Psychological testing on the Internet: new problems, old issues. *American Psychologist*, *59*(3), 150. https://doi.org/10.1037/0003-066X.59.3.150

Nantais, K. M., & Schellenberg, E. G. (1999). The Mozart effect: An artifact of preference. *Psychological Science*, *10*(4), 370-373. https://doi.org/10.1111/1467-9280.00170

Nieminen, S., Istók, E., Brattico, E., Tervaniemi, M., & Huotilainen, M. (2011). The development of aesthetic responses to music and their underlying neural and psychological mechanisms. *Cortex*: *A Journal Devoted to the Study of the Nervous System and Behavior, 47*(9), 1138-1146. https://doi.org/10.1016/j.cortex.2011.05.008

Novelli, G., Papagno, C., Capitani, E., Laiacona, M. (1986). Tre test clinici di memoria verbale a lungo termine: taratura su soggetti normali. *Archivio di Psicologia, Neurologia e Psichiatria*, *47*, 278–296.

Nussenbaum, K., Scheuplein, M., Phaneuf, C. M., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra: Psychology*, *6*(1), 17213. https://doi.org/10.1525/collabra.17213

Olderbak, S., Wilhelm, O. (2017). Emotion perception and empathy: An individual differences test of relations. *Emotion*, *17*(7), 1092–1106. https://doi.org/10.1037/emo0000308

Padulo, C., Mammarella, N., Brancucci, A., Altamura, M., & Fairfield, B. (2020). The effects of music on spatial reasoning. *Psychological Research*, *84*(6), 1723-1728. https://doi.org/10.1007/s00426-019-01182-6

Palomar-García, M. Á., Olcina-Sempere, G., Hernández Pardo, M., Parcet, M. A., Mandell, J. C., & Ávila, C. (2020). The Jake Mendell Test as a measure of individual differences in pitch discrimination: validity and reliability properties. *RECIEM, 17, 143-151*.

Palumbo, R., Mammarella, N., Di Domenico, A., & Fairfield, B. (2018). When and where in aging: the role of music on source monitoring. *Aging Clinical and Experimental Research*, *30*(6), 669-676.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, *30*, 653-661. https://doi.org/10.1097/AUD.0b013e3181b412e9

Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, *49*(12), 3338-3345. https://doi.org/10.1016/j.neuropsychologia.2011.08.007

Park, H. L., O'Connell, J. E., & Thomson, R. G. (2003). A systematic review of cognitive decline in the general elderly population. *International journal of geriatric psychiatry*, *18*(12), 1121-1134. https://doi.org/10.1002/gps.1023

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, *2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Pearce, M. T., & Halpern, A. R. (2015). Age-related patterns in emotions evoked by music. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(3), 248–253. https://doi.org/10.1037/a0039279

Penhune, V. B., & de Villers-Sidani, E. (2014). Time for new thinking about sensitive periods. *Frontiers in Systems Neuroscience*, *8*, 55. https://doi.org/10.3389/fnsys.2014.00055

Penhune, V. B. (2019). Musical expertise and brain structure: The causes and consequences of training. In M. H. Thaut & D. A. Hedges, (Eds.), *The Oxford handbook of music and the brain* (pp. 417– 438). Oxford University Press. https://doi.org/10.1093/oxfordhb/ 9780198804123.013.17

Penhune, V. B. (2020). A gene-maturation-environment model for un- derstanding sensitive period effects in musical training. *Current Opinion in Behavioral Sciences*, *36*, 13–22. https://doi.org/10. 1016/j.cobeha.2020.05.011

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. *Annals of the New York Academy of Sciences*, *999*, 58– 75. https://doi.org/10.1196/annals.1284.006

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, *68*(2), 111-141.

Phillips, L. H., Smith, L., & Gilhooly, K. J. (2002). The effects of adult aging and induced positive and negative mood on planning. *Emotion*, *2*(3), 263–272. https://doi.org/10.1037/1528-3542.2.3.263

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Mozart effect–Shmozart effect: A meta-analysis. *Intelligence*, *38*(3), 314-323. https://doi.org/10.1016/J.INTELL.2010.03.001

Puryear, J. S., Kettler, T., & Rinn, A. N. (2017). Relationships of personality to differential conceptions of creativity: A systematic review. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(1), 59-68. https://doi.org/10.1037/aca0000079

Rabinowitch, T. C. (2020). The potential of music to effect social change. *Music & Science*, *3*, https://doi.org/2059204320939772.

Radocy, R. E., & Boyle, J. D. (2012). (5th ed.). *Psychological foundations of musical behavior* Charles C Thomas Publisher.

Rauscher, F. H., Shaw, G. L., & Ky, C. N. (1993). Music and spatial task performance. *Nature*, *365*(6447), 611. https://doi.org/10.1038/365611a0

Raven, J. C. (1965). Advance Progressive Matrices, Sets I and II. Psychological Corporation.

Reardon MacLellan, C. (2011). Differences in Myers-Briggs personality types among high school band, orchestra, and choir members. *Journal of research in music education*, *59*(1), 85-100.

Reaves, S., Graham, B., Grahn, J., Rabannifard, P., & Duarte, A. (2016). Turn off the music! Music impairs visual associative memory performance in older adults. *Gerontologist*, *56*(3), 569-577. https://doi.org/10.1093/geront/gnu113

Rose, D., Bartoli, A. J., & Heaton, P. (2019). Formal-informal musical learning, sex and musicians' personalities. *Personality and Individual Differences*, *142*, 207-213. https://doi.org/10.1016/j.paid.2018.07.015

Ross-Sheehy, S., Reynolds, E., & Eschman, B. (2021). Unsupervised online assessment of visual working memory in 4-to 10-year-old children: array size influences capacity estimates and task performance. *Frontiers in Psychology*, *12*(692228). https://doi.org/10.3389/fpsyg.2021.692228

Sala, G., & Gobet, F. (2020). Cognitive and academic benefits of music training with children: A multilevel meta-analysis. *Memory & Cognition*, *48*(8), 1429–1441. https://doi.org/10.3758/s13421-020-01060-2

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*(4), 507–514. https://doi.org/10.1016/j.neurobiolaging.2008.09.023

Sauce, B., & Matzel, L. D. (2018). The paradox of intelligence: Heritability and malleability coexist in hidden gene-environment interplay. *Psychological Bulletin*, *144*(1), 26–47. https://doi.org/10.1037/bul0000131

Schaal, N. K., Banissy, M. J., & Lange, K. (2015). The rhythm span task: Comparing memory capacity for musical rhythms in musicians and non-musicians. *Journal of New Music Research*, *44*(1), 3–10. https://doi.org/10.1080/09298215.2014.937724

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology*, *98*(2), 457– 468. https://doi.org/10.1037/0022-0663.98.2.457

Schellenberg, E. G. (2011). Music lessons, emotional intelligence, and IQ. *Music Perception, 29*, 185-194. https://doi.org/10.1525/mp.2011.29.2.185

Schellenberg, E. G. (2019). Music training, music aptitude, and speech perception. *Proceedings of the National Academy of Sciences*, *116*(8), 2783-2784. https://doi.org/10.1073/pnas.1821109116

Schellenberg, E. G. (2020). Correlation = causation? Music training, psychology, and neuroscience. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(4), 475. https://doi.org/10.1037/aca0000263

Schellenberg, E. G., & Weiss, M. W. (2013). Music and cognitive abilities. In D. Deutsch (Ed.), *The psychology of music* (pp. 499–550). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-381460-9.00012-2

Schellenberg, E. G., & Mankarious, M. (2012). Music training and emotion comprehension in childhood. *Emotion*, *12*(5), 887–891. https:// doi.org/10.1037/a0027971

Schellenberg, E. G., & Hallam, S. (2005). Music listening and cognitive abilities in 10-and 11-year-olds: The blur effect. *Annals of the New York Academy of Sciences*, *1060*(1), 202-209. https://doi.org/10.1196/annals.1360.013

Schellenberg, E. G., Nakata, T., Hunter, P. G., & Tamoto, S. (2007). Exposure to music and cognitive performance: tests of children and adults. *Psychology of Music*, *35*, 5–19. https://doi.org/10.1177/0305735607068885

Schellenberg, E. G. (2012). Cognitive performance after listening to music: A review of the Mozart effect. In R. A. R. MacDonald, G. Kreutz, & L. Mitchell (Eds.), Music, health, and wellbeing (pp. 324–338). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586974.003.0022

Seashore, C. (1919). *The psychology of musical talent*. Holt.

Seashore, C. E., Saetveit, J. G., & Lewis, D. (1960). *The Seashore measures of musical talent* (rev. ed.). Psychological Corporation.

Shuter-Dyson, R. (2000). Profiling music students: Personality and religiosity. *Psychology of Music*, *28*(2), 190-196. https://doi.org/10.1177/0305735600282008

Silva, S., Belim, F., & Castro, S. L. (2020). The Mozart effect on the episodic memory of healthy adults is null, but low-functioning older adults may be an exception. *Frontiers in Psychology*, *11*, 538194. https://doi.org/10.3389/fpsyg.2020.538194

Simonton, D. K. (2006). Presidential IQ, openness, intellectual brilliance, and leadership: Estimates and correlations for 42 U.S. chief executives. *Political Psychology*, *27*(4), 511–639. https://doi.org/10.1111/j.1467 -9221.2006.00524.x

Simonton, D. K. (2009). Creative genius in classical music. *The Psychologist*, *22*(12), 1076–1078.

Sloboda, J. A. (1992). Empirical studies of emotional response to music. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication* (pp. 33–46). American Psychological Association. https://doi.org/10.1037/10104-003

Steele, C. J., & Zatorre, R. J. (2018). Practice makes plasticity. *Nature Neuroscience*, *21*(12), 1645–1646. https://doi.org/10.1038/s41593-018-0280-4

Stern, Y., Zarahn, E., Hilton, H. J., Flynn, J., DeLaPaz, R., & Rakitin, B. (2003). Exploring the neural basis of cognitive reserve. *Journal of clinical and experimental neuropsychology*, *25*(5), 691-701. https://doi.org/10.1076/jcen.25.5.691.14573

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of intelligence*. Cambridge University Press.

Strait, D., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*, 22-29. https://doi.org/10.1016/j.heares.2009.12.021

Sutcliffe, R., Rendell, P. G., Henry, J. D., Bailey, P. E., Ruffman, T. (2017). Music to my ears: Age-related decline in musical and facial emotion recognition. *Psychology of Aging*, *32*(8), 698–709. https://doi.org/10.1037/pag0000203

Swaminathan, S., Kragness, H. E., & Schellenberg, E. G. (2021). The Musical Ear Test: Norms and correlates from a large sample of Canadian undergraduates. *Behavior Research Methods*. Advance online publication. https://doi.org/10.3758/s13428-020-01528-8

Swaminathan, S., & Schellenberg, E. G. (2019). Music training and cognitive abilities: Associations, causes, and consequences. In *The Oxford handbook of music and the brain*, 644-670.

Swaminathan, S., & Schellenberg, E. G. (2020). Musical ability, music training, and language ability in childhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(12), 2340–2348. https://doi.org/10.1037/xlm0000798

Swaminathan, S., Schellenberg, E. G., & Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence*, *62*, 119–124. https://doi.org/10.1016/j .intell.2017.03.005

Swaminathan, S., & Schellenberg, E. G. (2018). Musical competence is predicted by music training, cognitive abilities, and personality. *Scientific Reports*, *8*(1), 9223. https://doi.org/10.1038/s41598-018-27571-2

Swaminathan, S., & Schellenberg, E. G. (2021). Music training. In T. Strobach, J. Karbach (Eds.), *Cognitive training* (pp. 307-318). Springer, Cham.

Sylvia, P. J., Nusbaum, E. C., Berg, C., Martin, C., & O'Connor, A. (2009). Openness to experience, plasticity, and creativity: Exploring lower-order, high-order, and interactive effects. *Journal of Research in Personality*, *43*(6), 1087–1090. https://doi.org/10.1016/j.jrp.2009.04.015

Taherbhai, H., Seo, D., & Bowman, T. (2012). Comparison of paper– pencil and online performances of students with learning disabilities. *British Educational Research Journal*, *38*(1), 61–74. https://doi.org/ 10.1080/01411926.2010.526193

Talbot, M. (1971). Albinoni: The Professional Dilettante. *The Musical Times*, *112*(1540), 538–541. https://doi.org/10.2307/957427

Thomas, K. S., Silvia, P. J., Nusbaum, E. C., Beaty, R. E., & Hodges, D. A. (2016). Openness to experience and auditory discrimination ability in music: An investment approach. *Psychology of Music*, *44*(4), 792–801. https://doi.org/10.1177/0305735615592013

Thompson, R. G., Moulin, C. J. A., Hayre, S., & Jones, R. W. (2005). Music enhances category fluency in healthy older adults and Alzheimer's disease patients. *Experimental Aging Research.*, *31*, 91–99. https://doi.org/10.1080/03610730590882819

Thompson, W. F., Schellenberg, E. G., & Letnic, A. K. (2012). Fast and loud background music disrupts reading comprehension. *Psychology of Music*, *40*(6), 700-708. https://doi.org/10.1177/0305735611400173

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*(3), 248-251. https://doi.org/10.1111/1467-9280.00345

Tierney, A., & Kraus, N. (2013). Music training for the development of reading skills. *Progress in Brain Research*, *207*, 209–241. https://doi.org/10.1016/B978-0-444-63327-9.00008-4

Triller, N., Eržen, D., Duh, Š., Primožič, M. P., & Košnik, M. (2006). Music during bronchoscopic examination: the physiological effects. *Respiration*, *73*(1), 95-99. https://doi.org/10.1159/000089818

Tsuji, S., Amso, D., Cusack, R., Kirkham, N., & Oakes, L. M. (2022). Empirical Research at a Distance: New Methods for Developmental Science. *Frontiers in Psychology*, 3011.

Ullén, F., Mosing, M. A., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences*, *63*, 87–93. https://doi.org/10.1016/j. paid.2014.01.057

Vaag, J., Sund, E. R., & Bjerkeset, O. (2018). Five-factor personality profiles among Norwegian musicians compared to the general workforce. *Musicae Scientiae*, *22*(3), 434-445. https://doi.org/10.1177/1029864917709519

Västfjäll, D. (2001). Emotion induction through music: A review of the musical mood induction procedure. *Music Scientiae*, *5*(1 suppl), 173-211. https://doi.org/10.1177/10298649020050S107

Vieillard, S., Roy, M., & Peretz, I. (2012). Expressiveness in musical emotions. *Psychological Research*, *76*(5), 641-653. https://doi.org/10.1007/s00426-011-0361-4

von Krause, M., Radev, S.T., Voss, A. (2022). Mental speed is high until age 60 as revealed by analysis of over a million participants. *Nature Human Behaviour*, *6*, 700–708. https://doi.org/10.1038/s41562-021-01282-7

Wagenmakers, E., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H., Derks, K., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7.

Wagenmakers, E., Verhagen, J., & Ly, A. (2016). How to quantify the evidence for the absence of a correlation. *Behavior Research Methods*, *48*(2), 413–426. https://doi.org/10.3758/s13428-015-0593-0.

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

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). Corrigendum to "The Musical Ear Test, a new reliable test for measuring musical competence" [Learning and Individual Differences Volume 20 (3) (2010) 188–196]. *Learning and Individual Differences*, 20(6), 705. https://doi.org/10.1016/j.lindif. 2010.10.001

Wang, X., Ossher, L., & Reuter-Lorenz, P. A. (2015). Examining the relationship between skilled music training and attention. *Consciousness and Cognition*, *36*, 169–179. https://doi.org/10.1016/j.concog.2015.06.014

Ward, E. V., Isac, A., Donnelly, M., Van Puyvelde, M., & Franco, F. (2021). Memory improvement in aging as a function of exposure to mood-matching music. *Acta psychologica*, *212*, 103206.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *47*, 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

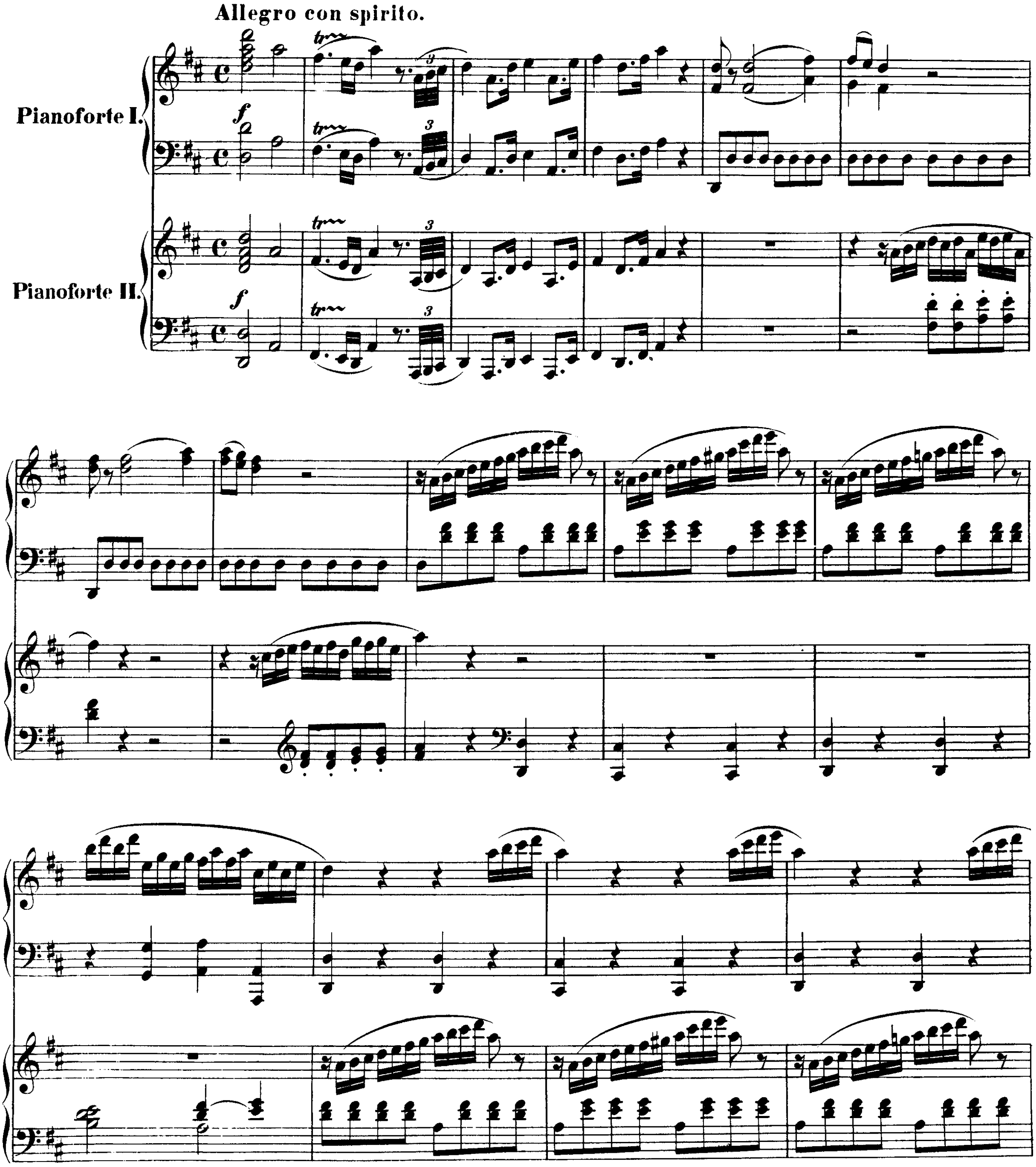
Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised (WAIS–R)*. Psychological Corporation.

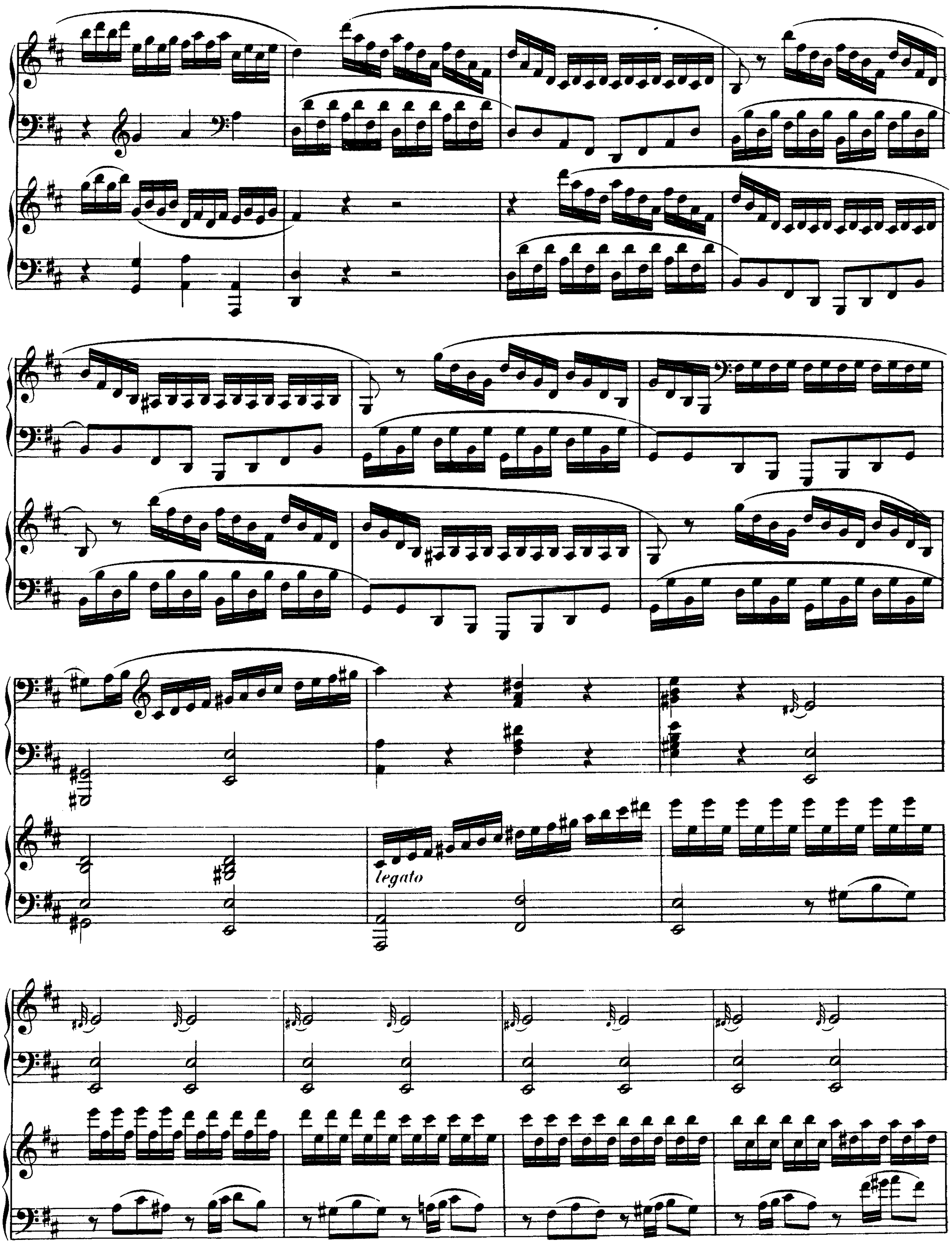
Wing, H. D. (1962). A revision of the Wing Musical Aptitude Test. *Journal of Research in Music Education*, *10*(1), 39–46. https://doi. org/10.2307/3343909

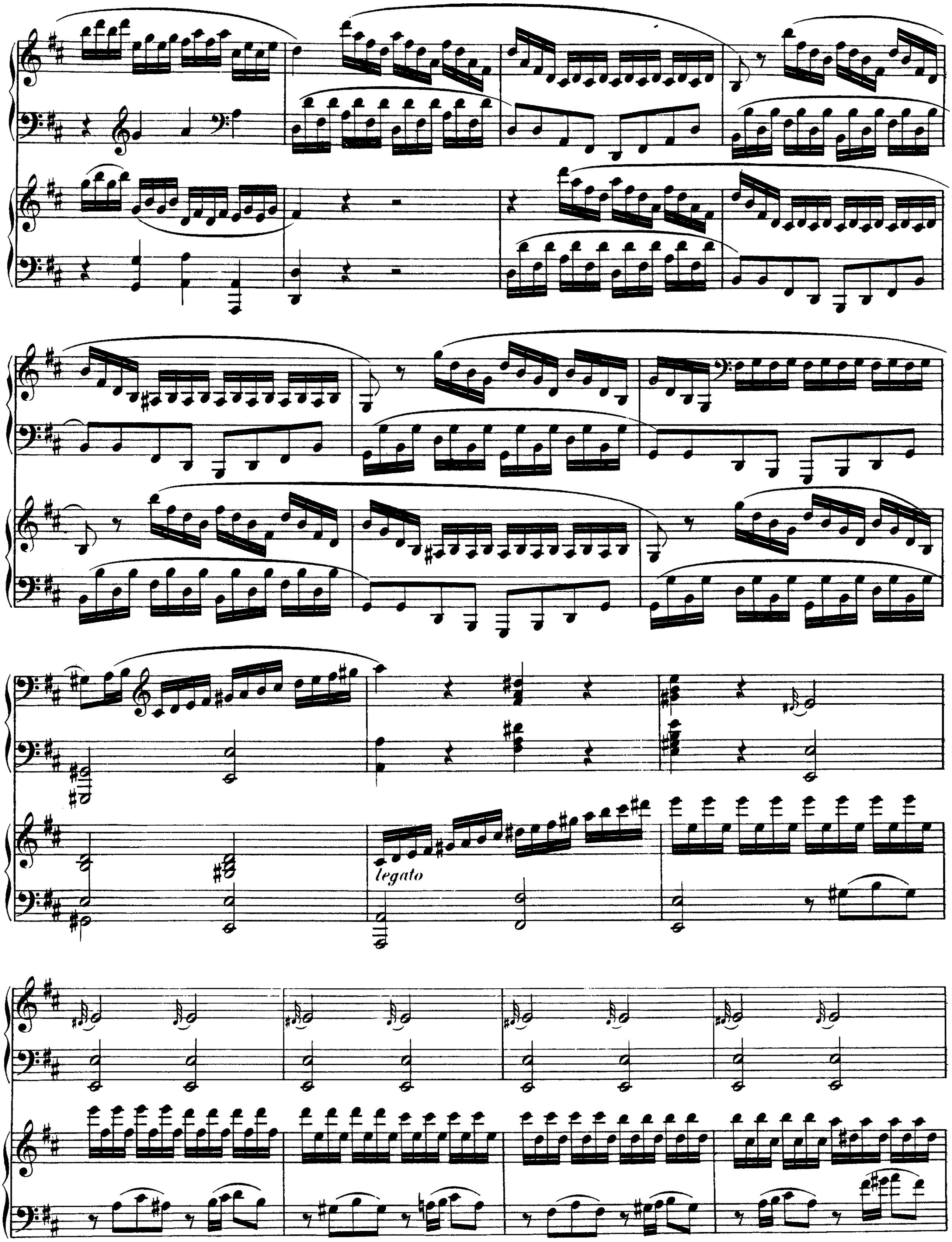
Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*, 565-585. https://doi.org/10.1017/S0142716407070312

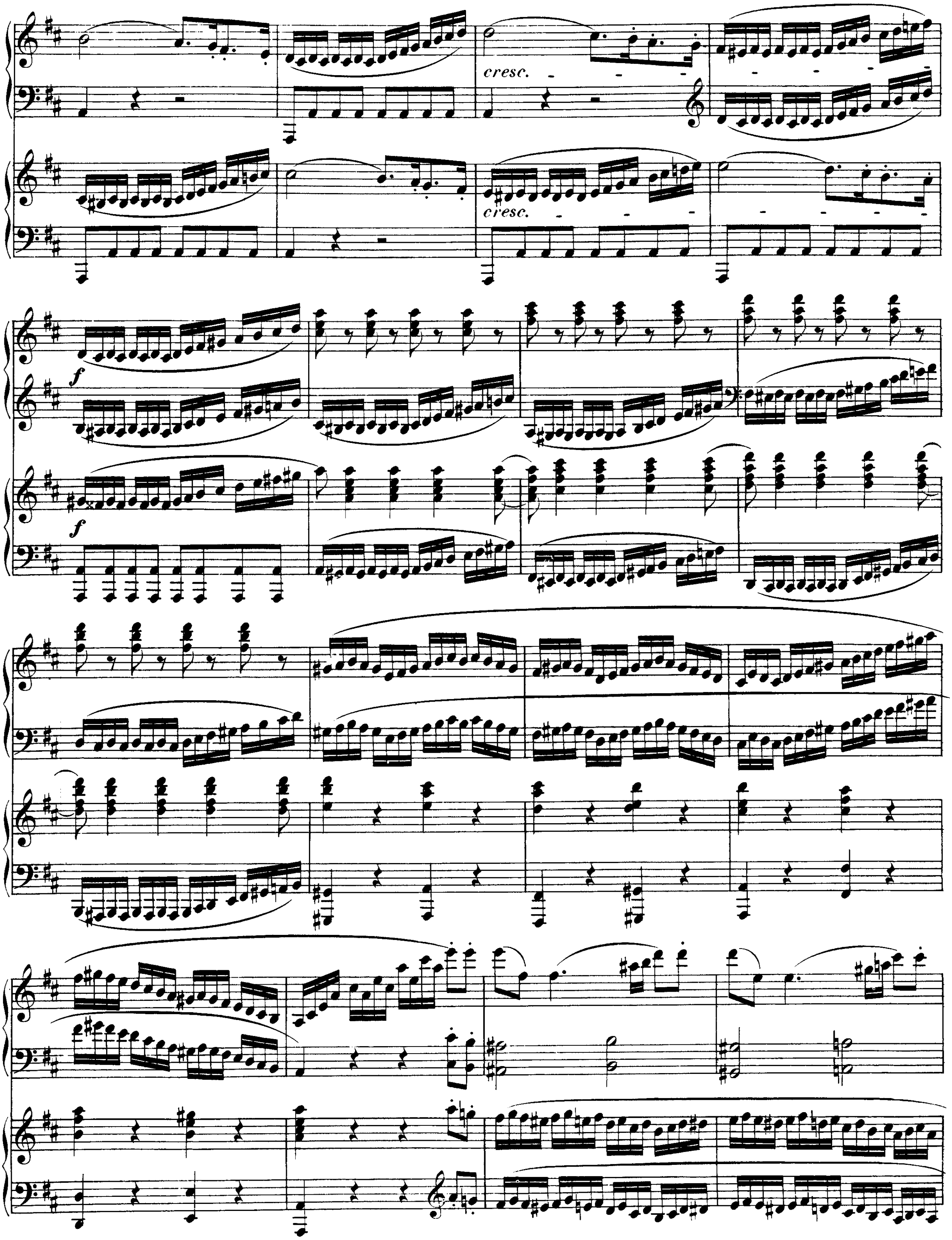
Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: development and validation of the Short-PROMS and the Mini-PROMS. *Annals of the New York Academy of Sciences*, *1400*(1), 33–45. https://doi.org/10.1111/nyas.13410

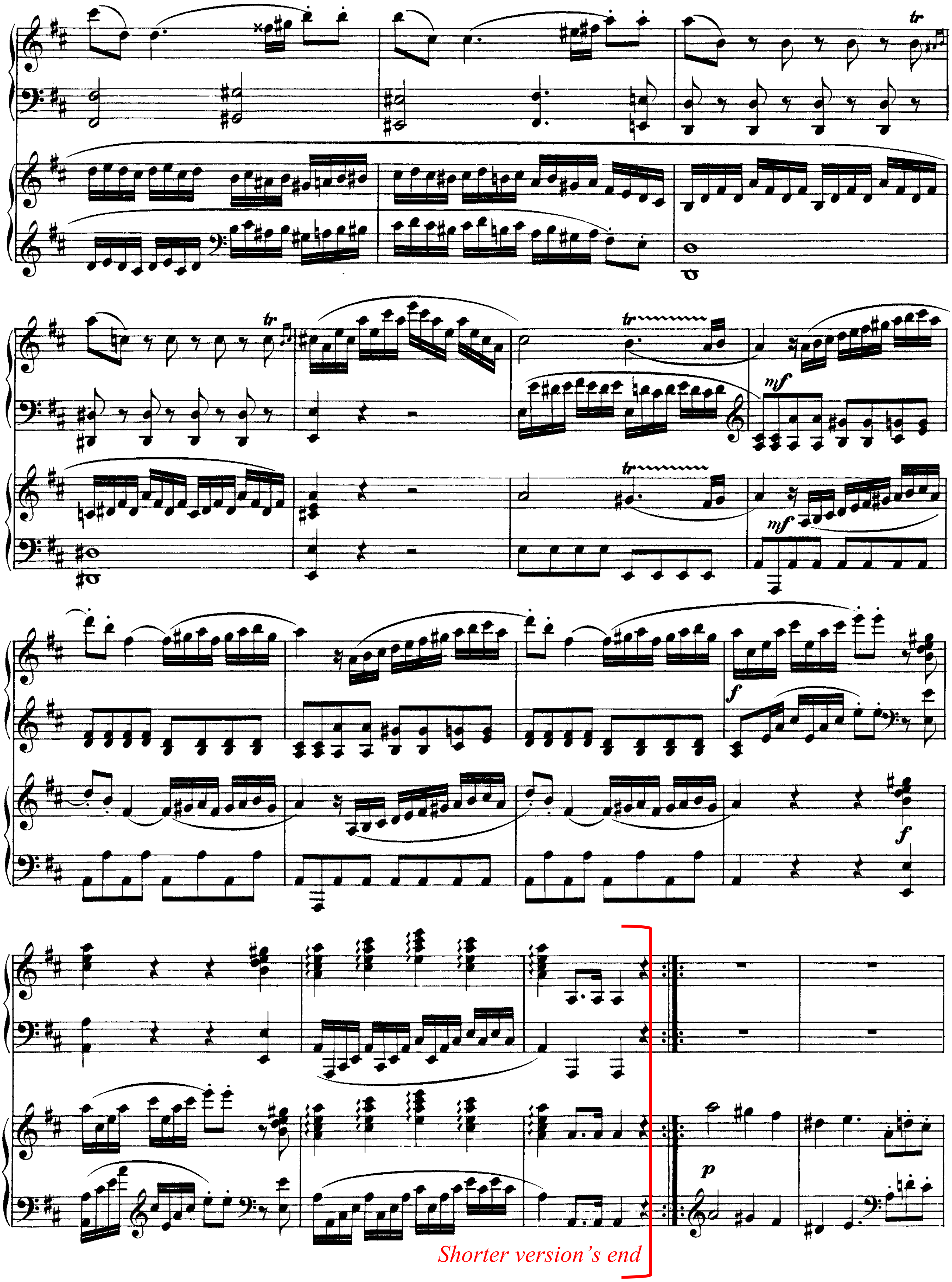
Zhang, J. D., Susino, M., McPherson, G. E., & Schubert, E. (2020). The definition of a musician in music psychology: A literature review and the six-year rule. *Psychology of Music*, *48*(3), 389–409. https://doi.org/ 10.1177/030573561880403

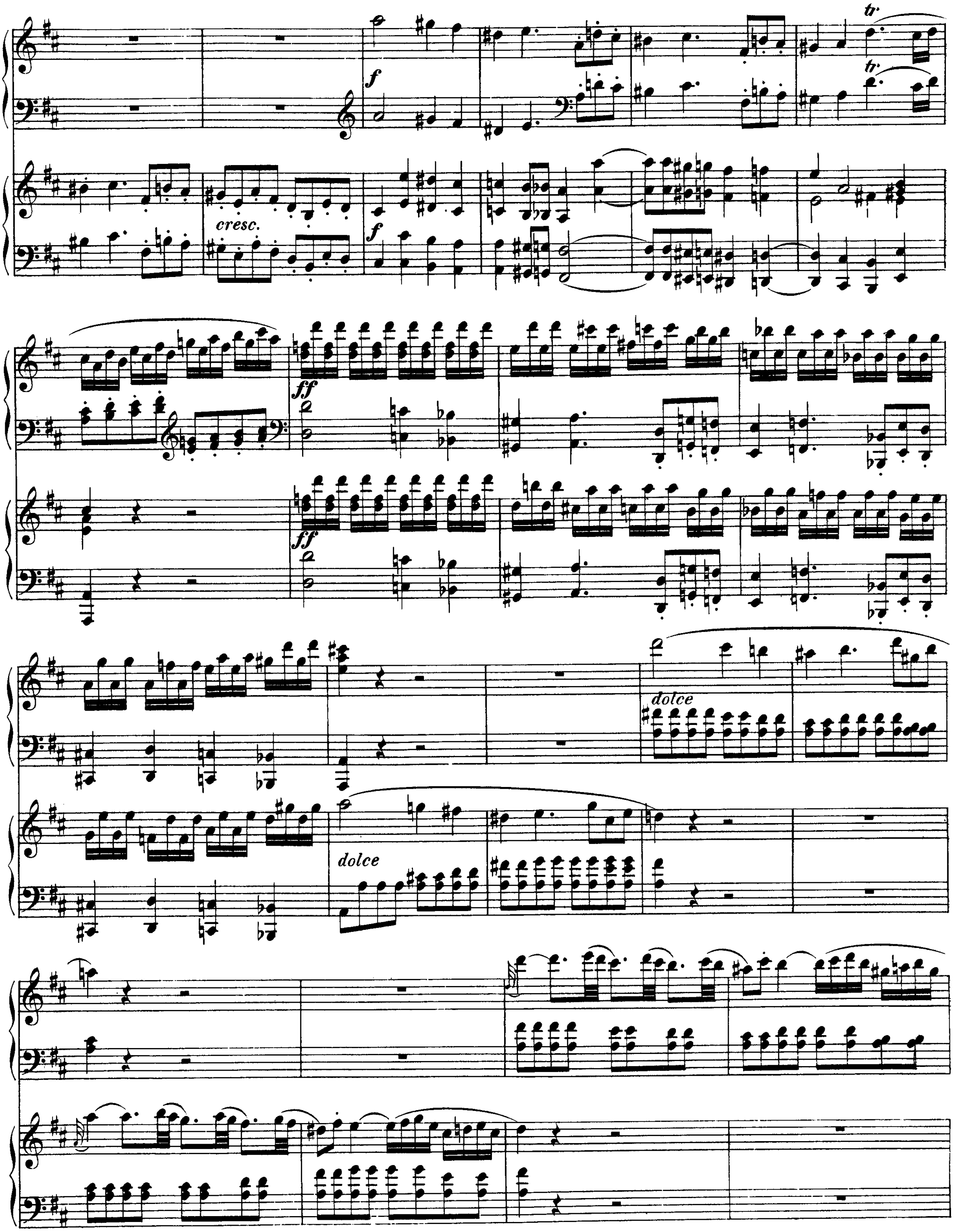
Appendix 1

---

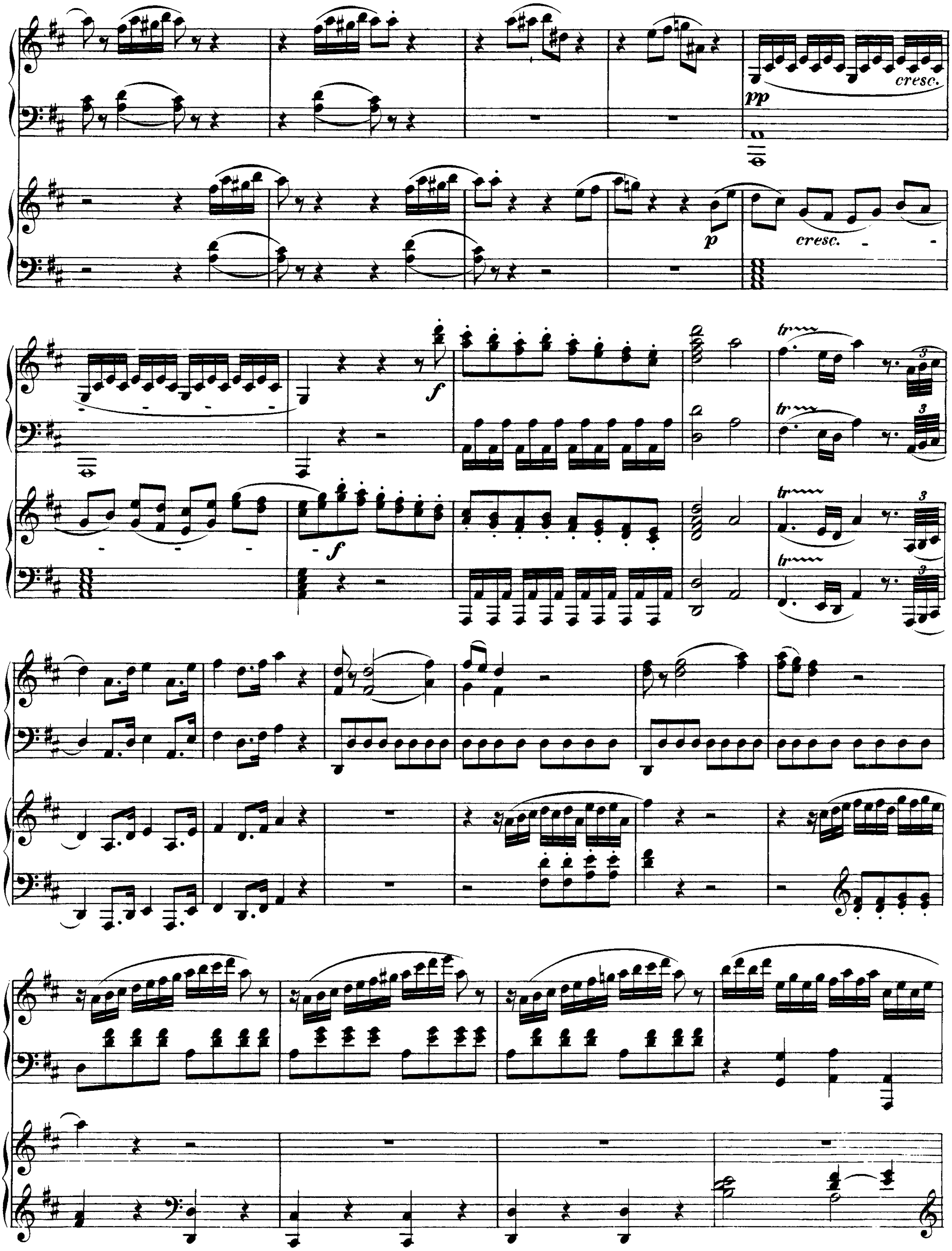
[8] First movement of the Mozart Sonata K 448 for two pianos. From *The Sheet Music Archive*, http://www.sheetmusicarchive.com.


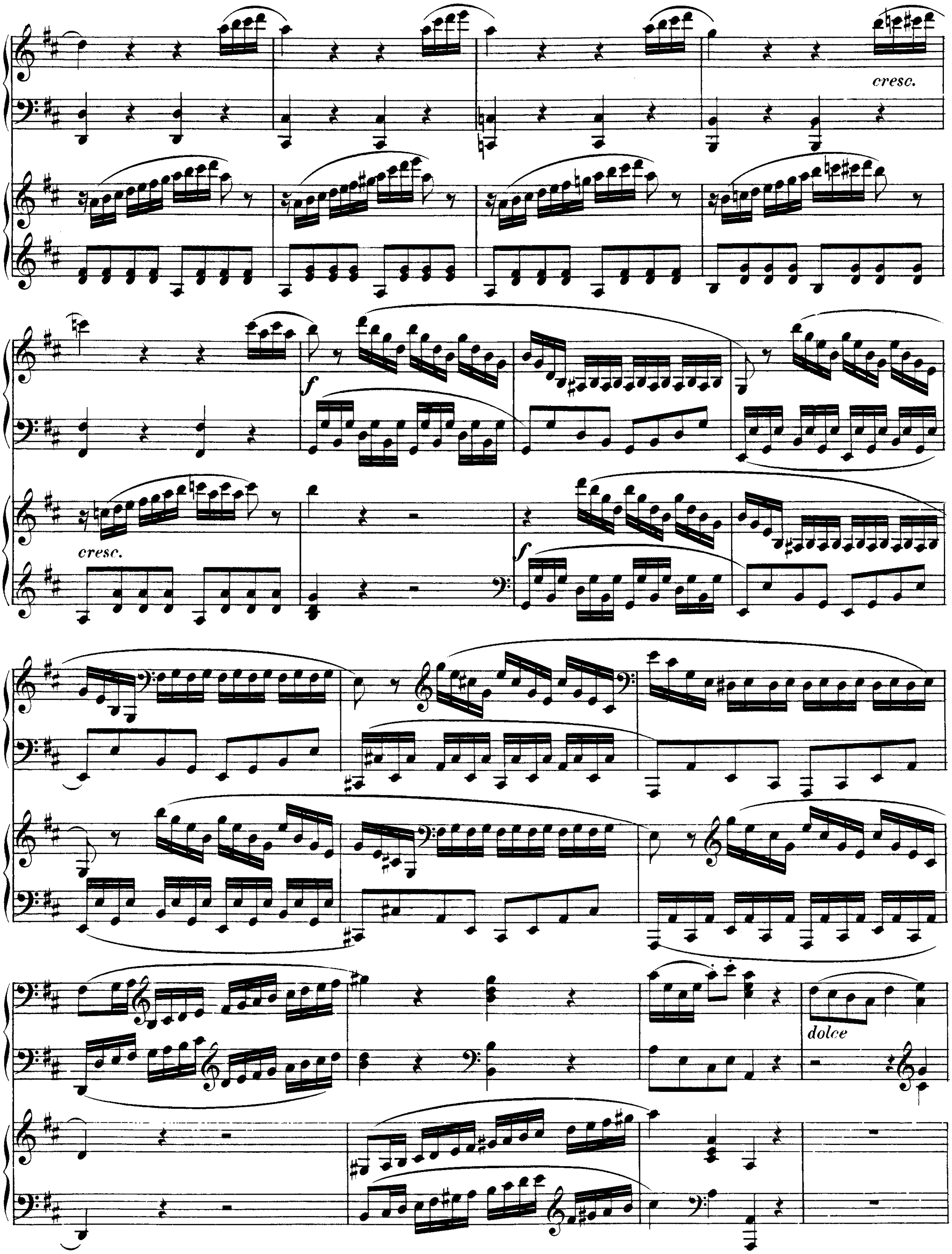
legato


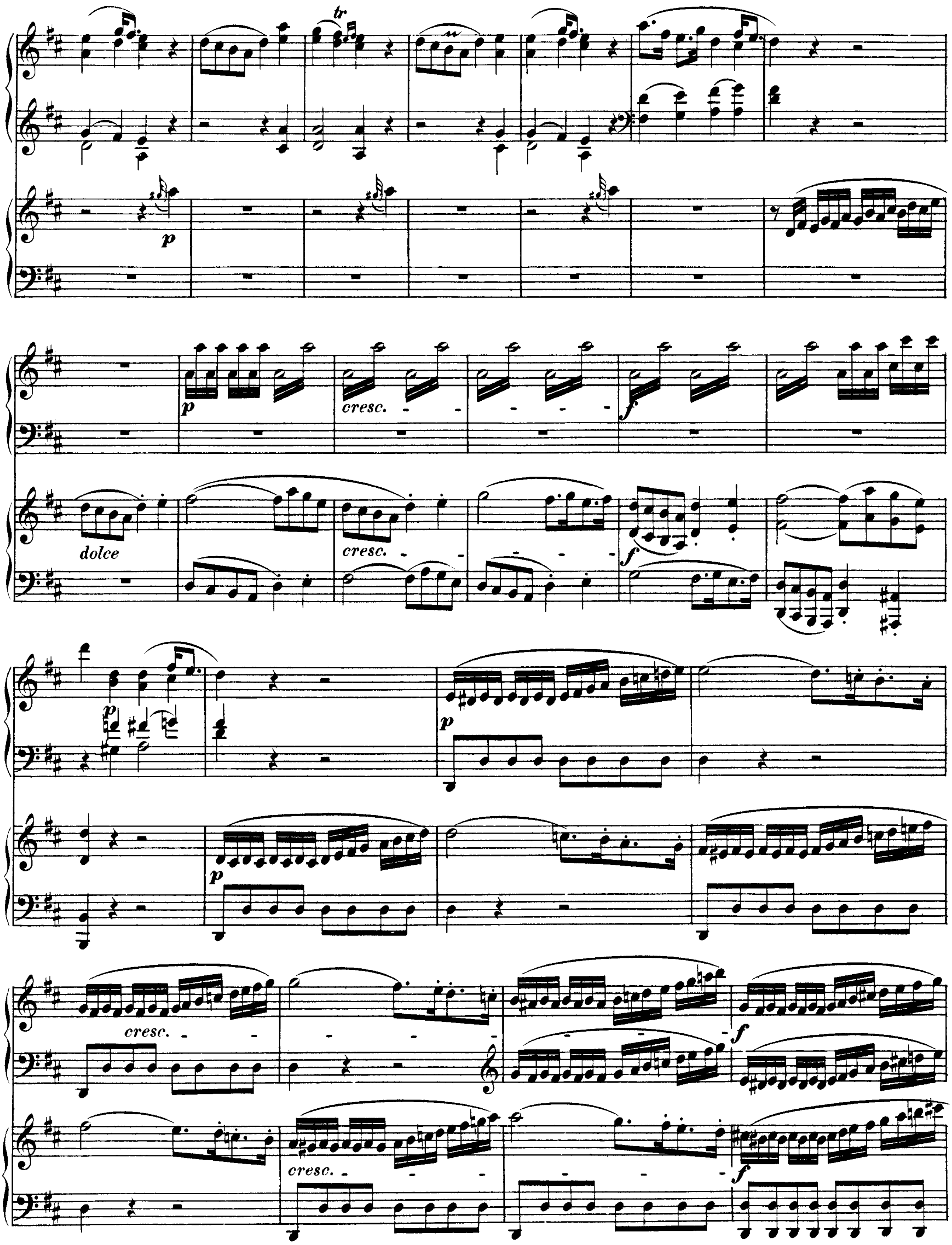
legato


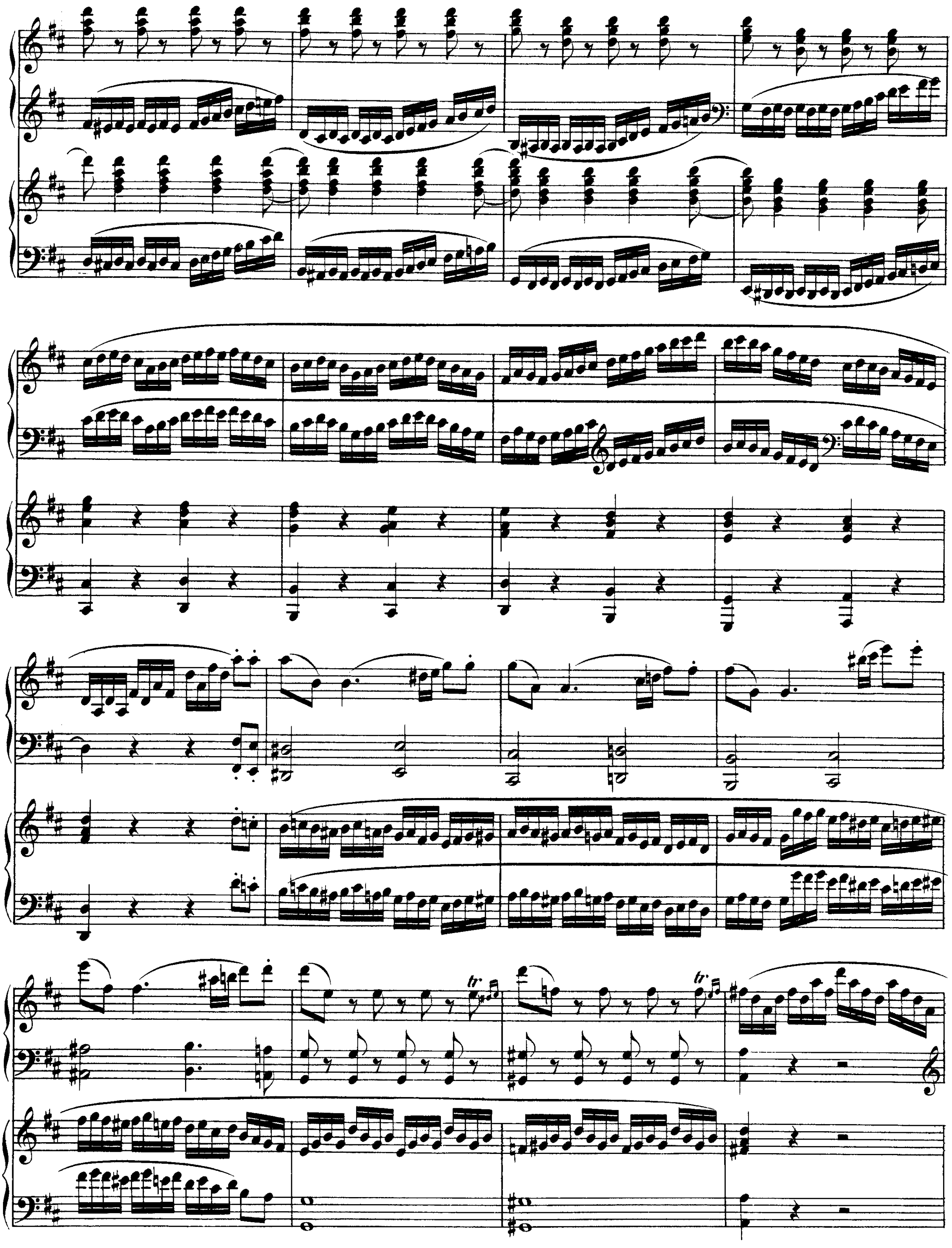
cresc.
cresc.
f
f


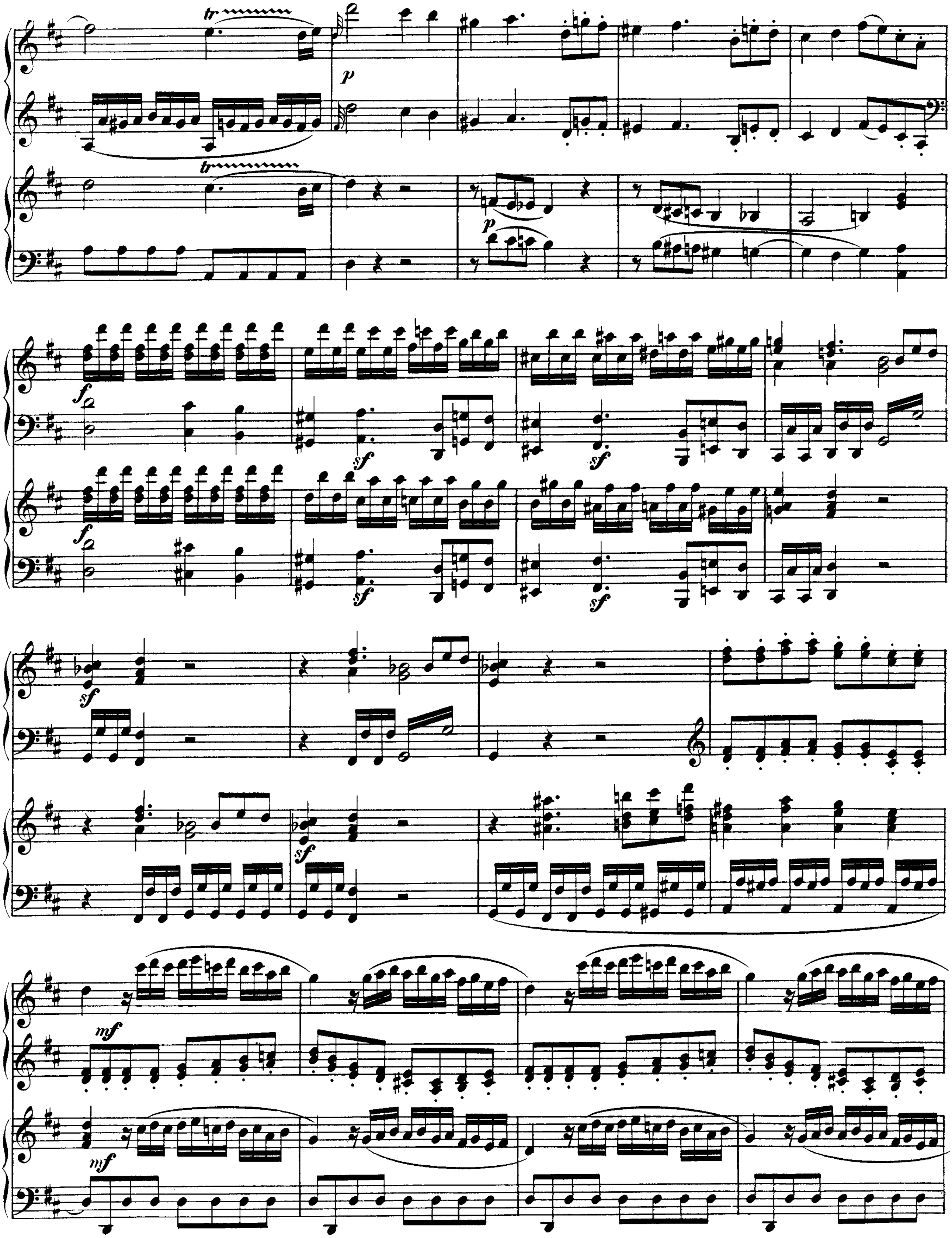
tr
mf
f
p
Shorter version's end


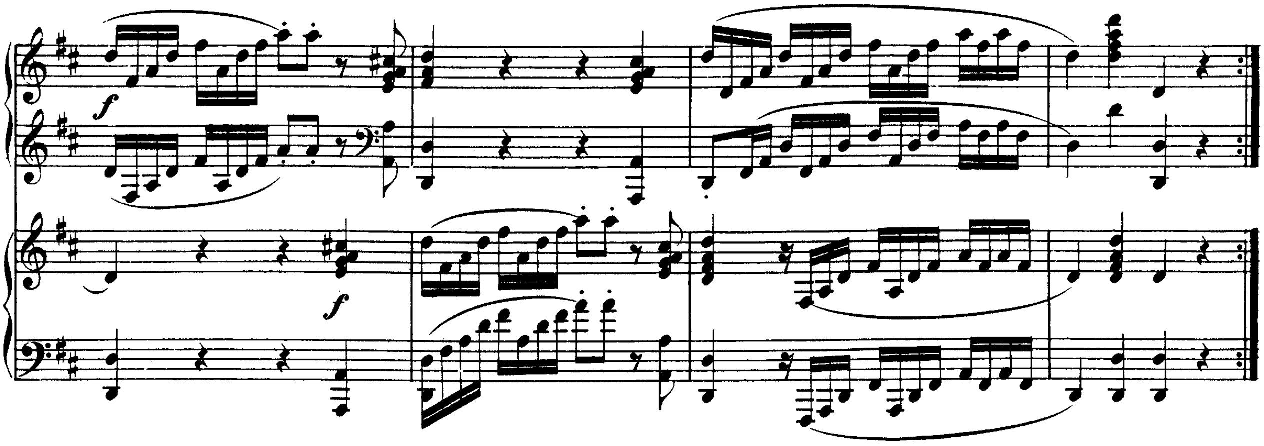
tr
tr
f
cresc.
f
ff
ff
dolce
dolce


pp
cresc.
p
cresc.
f
f
tr
3
tr
3
tr
3
tr
3


cresc.
cresc.
f
f
dolce


tr
p
p
cresc.
f
dolce
cresc.
f
p
p
p
cresc.
cresc.
f
f


tr
tr


tr
tr
p
p
f
f
sf
sf
sf
sf
sf
sf
mf
mf


f
f

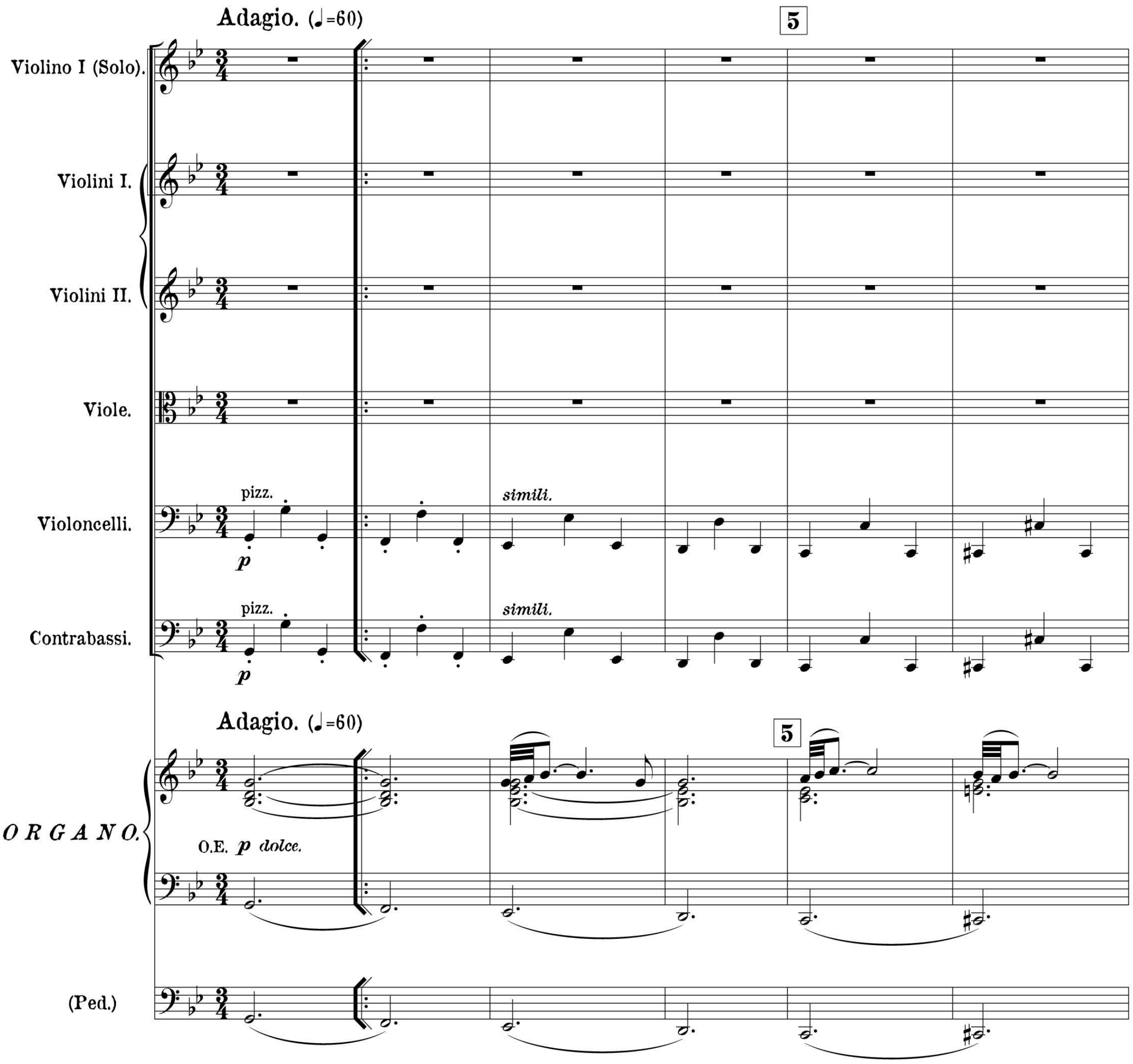
Appendix 2

Albinoni - Giazotto.

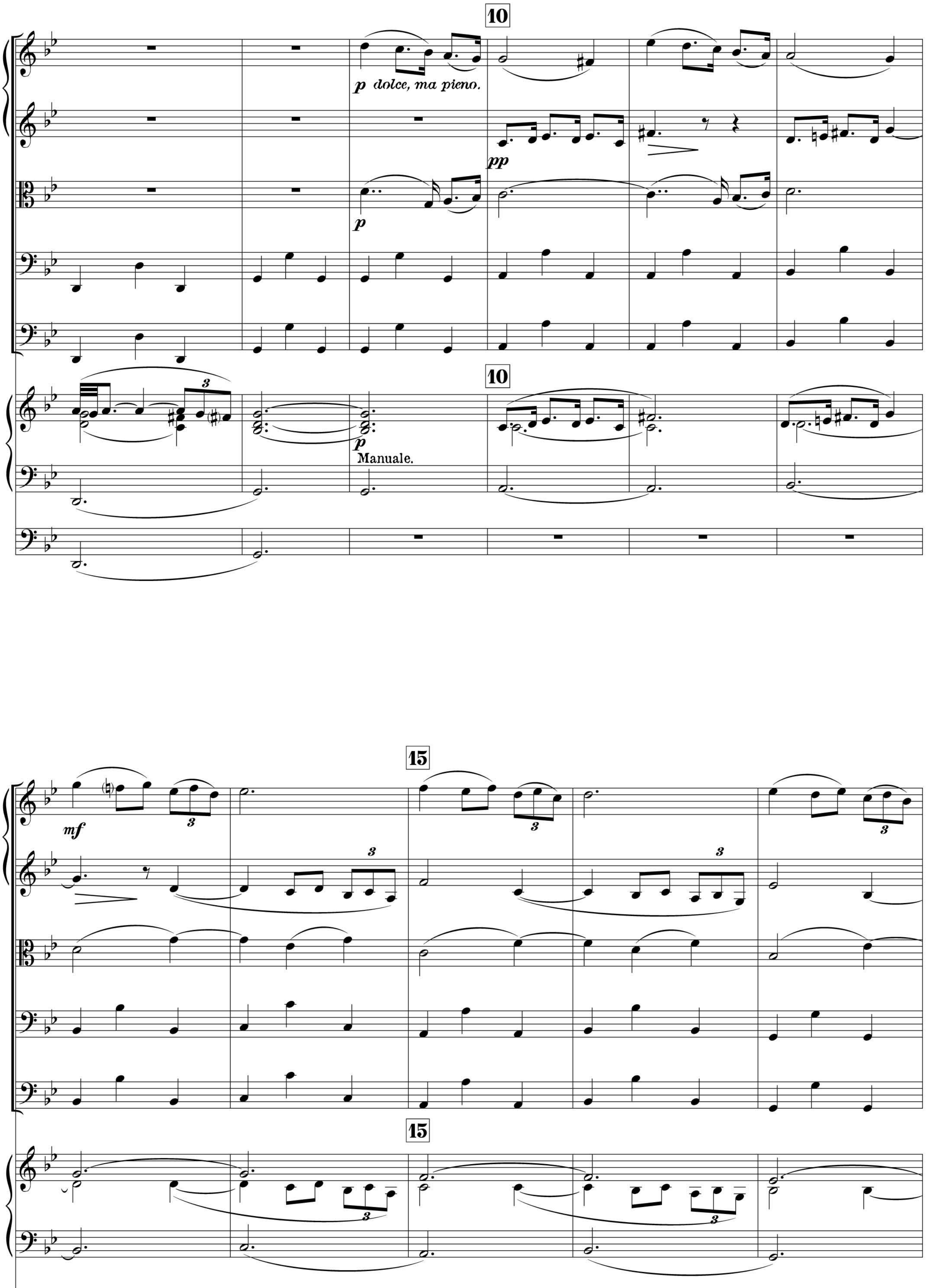
ADAGIO in SOL min.

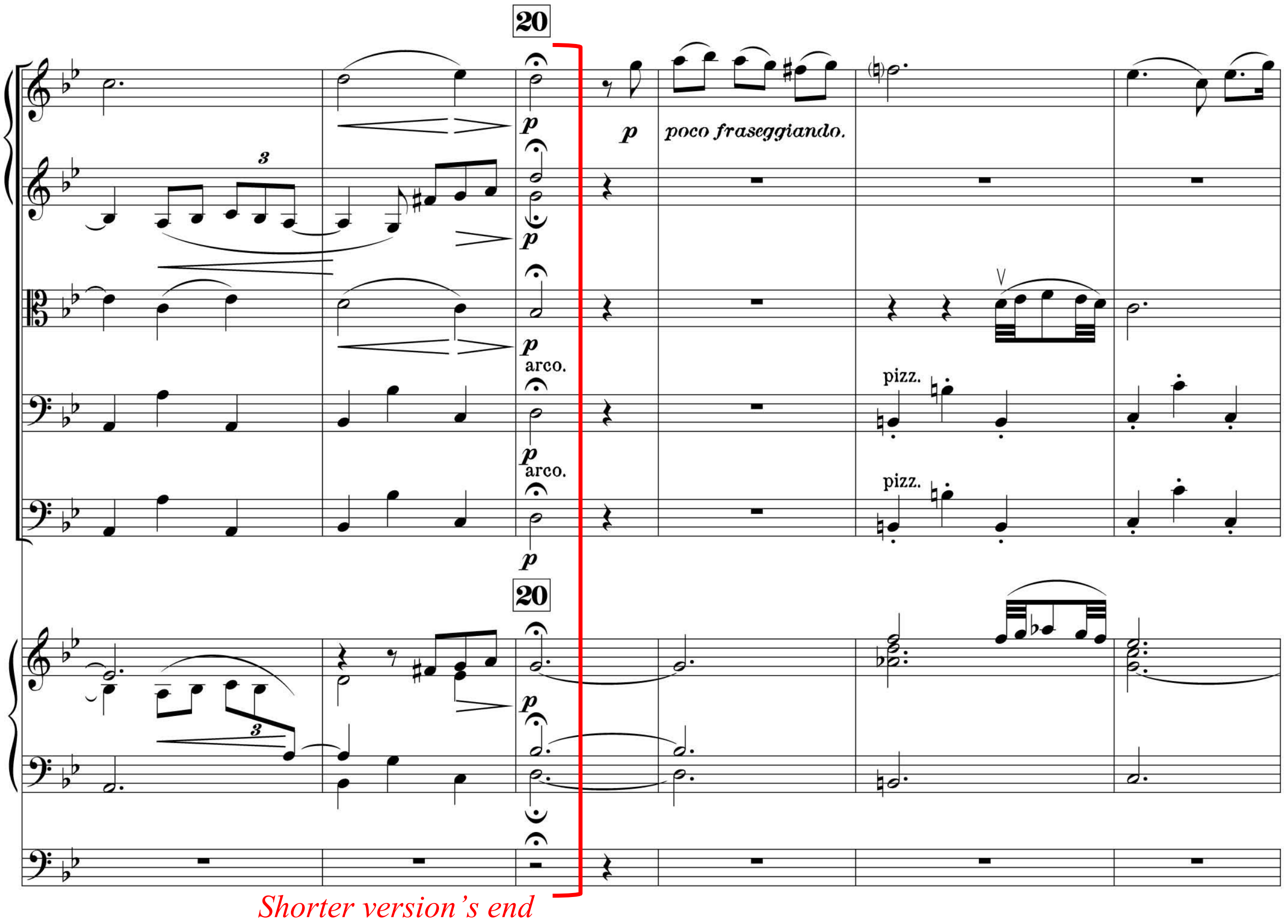
PER ARCHI E ORGANO

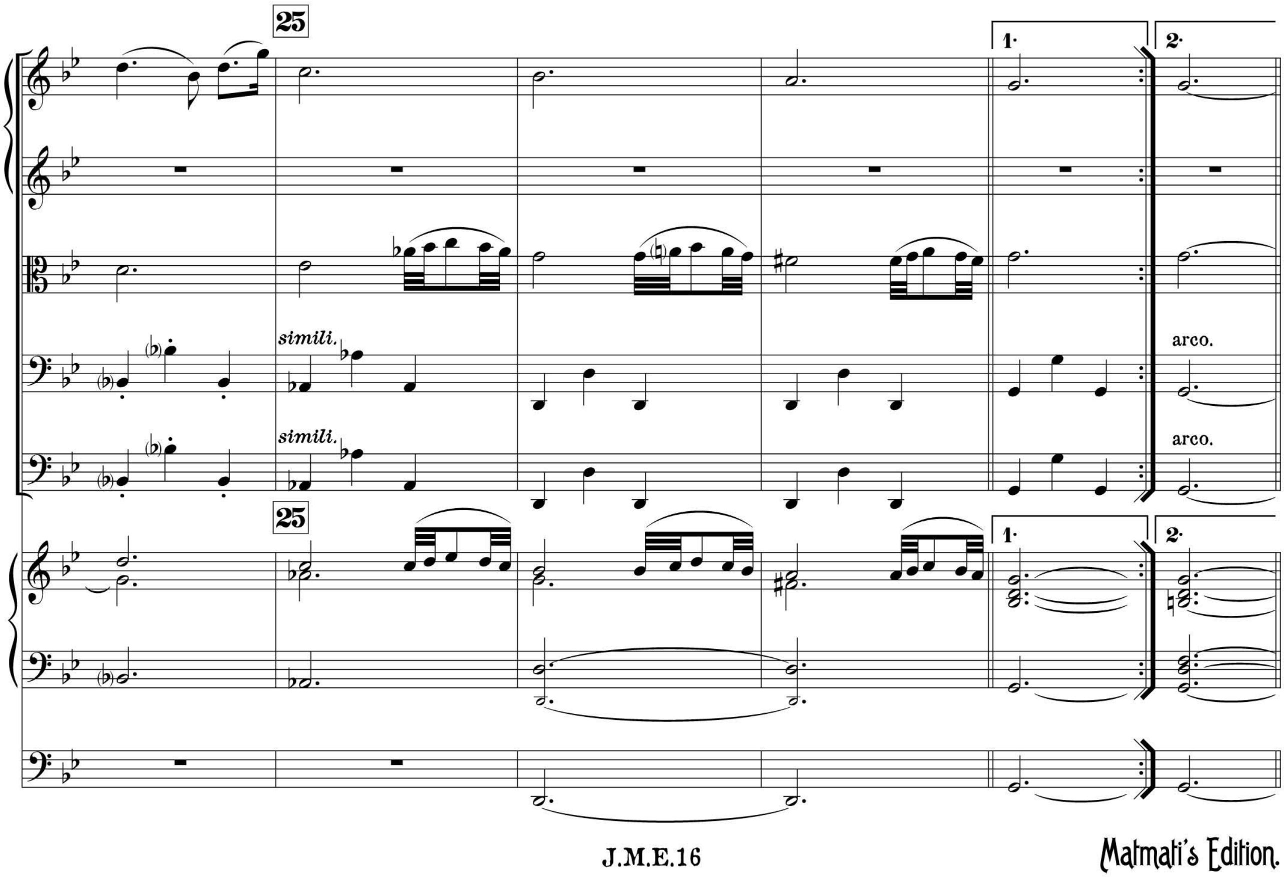
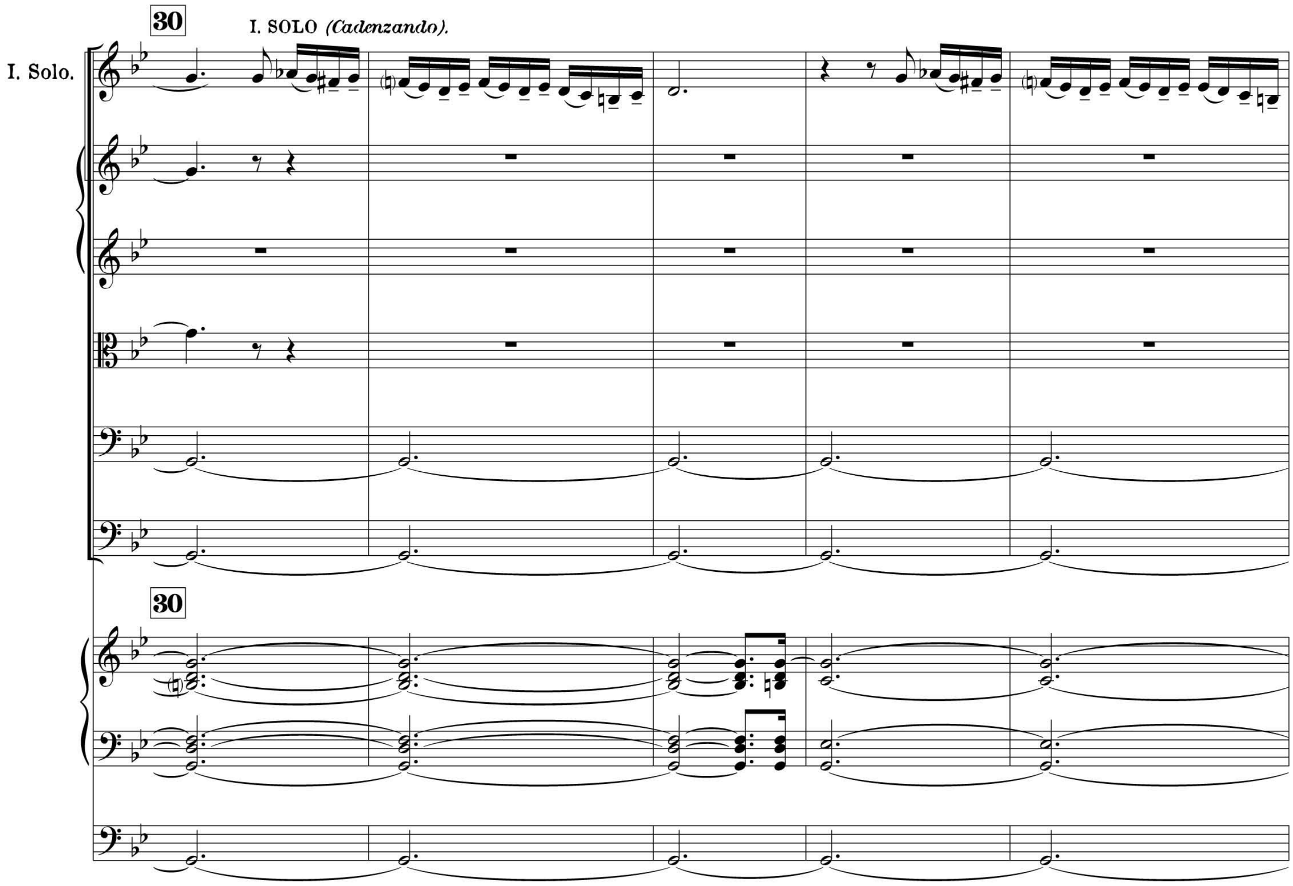
su un spunto tematico e su un basso numerato di Tomaso Albinoni

Neue Ausgabe von/ Edited by: Jawher Matmati
(1st September 2011)

[9]

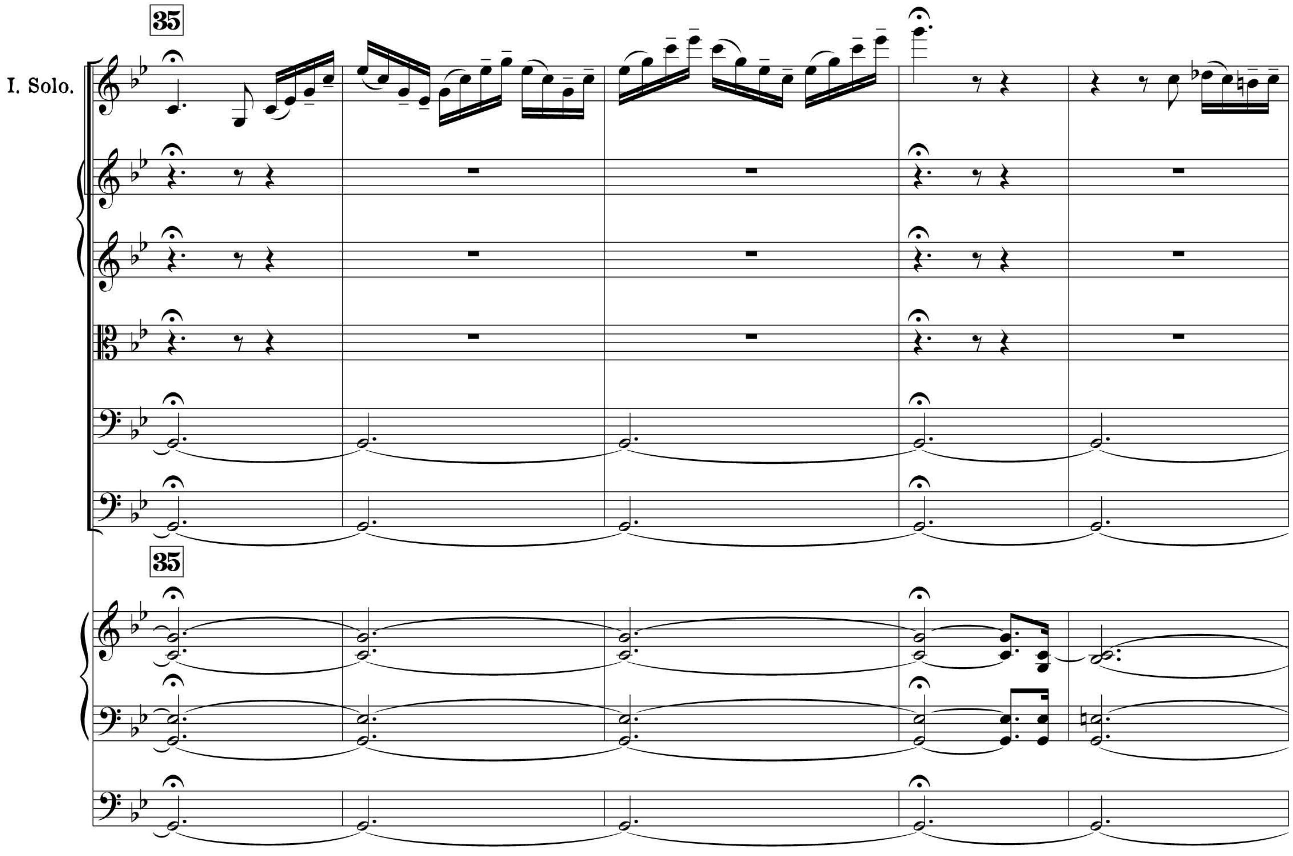
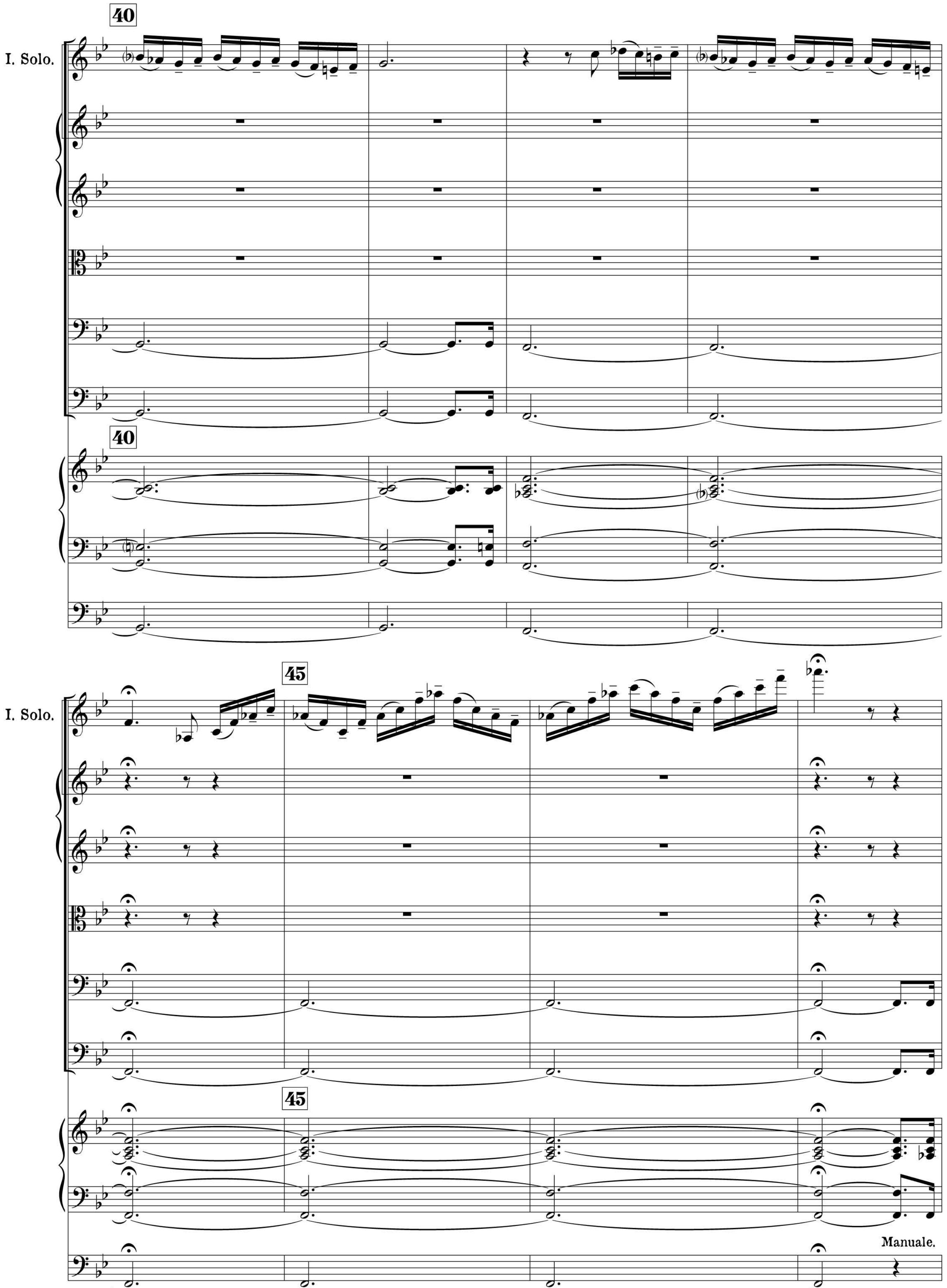
---

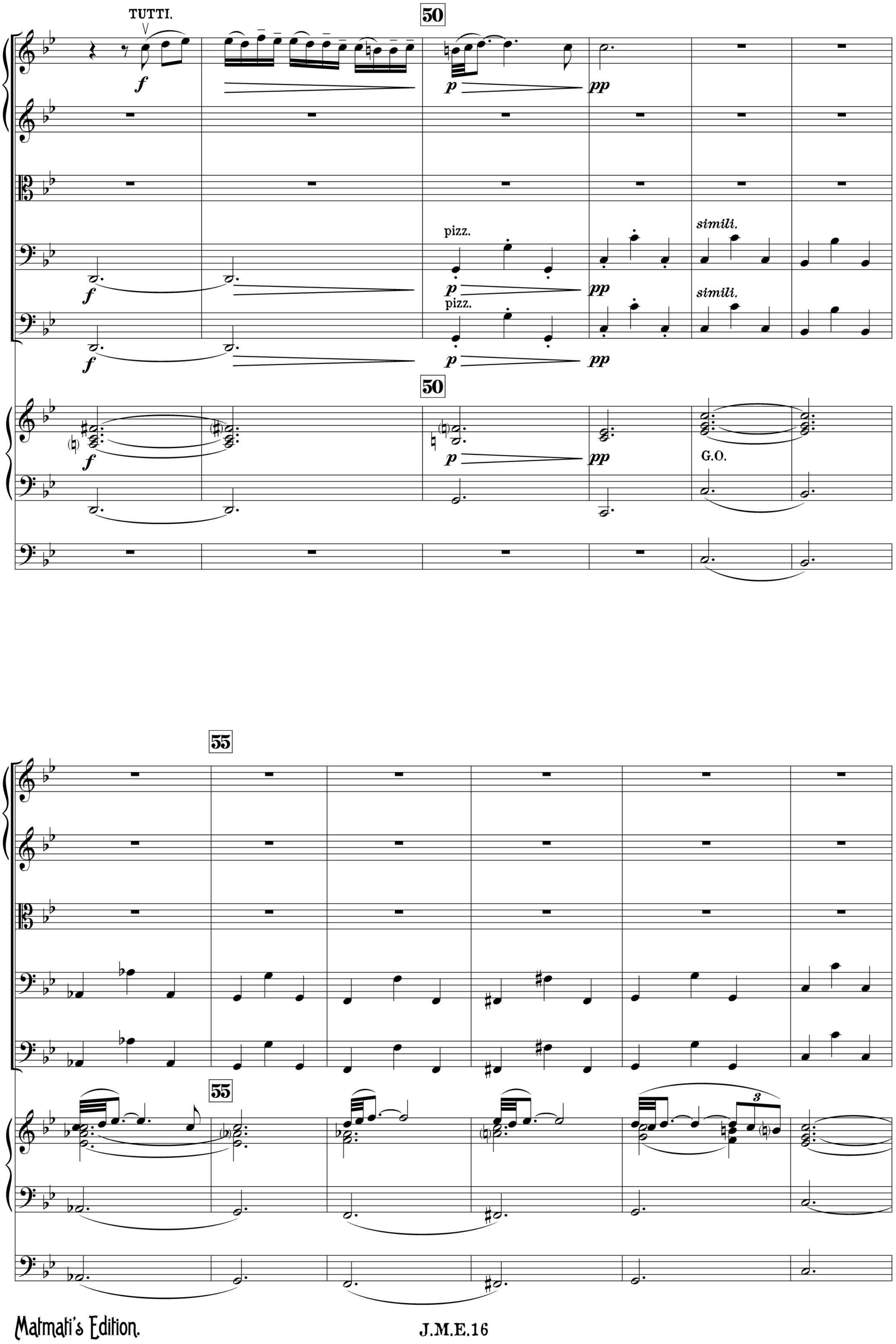
[9] Albinoni-Giazotto *Adagio* for organ and strings. From http://www.free-scores.com/Download-PDF-Sheet-Music-jojo1992.htm. This version is similar to the one we presented.


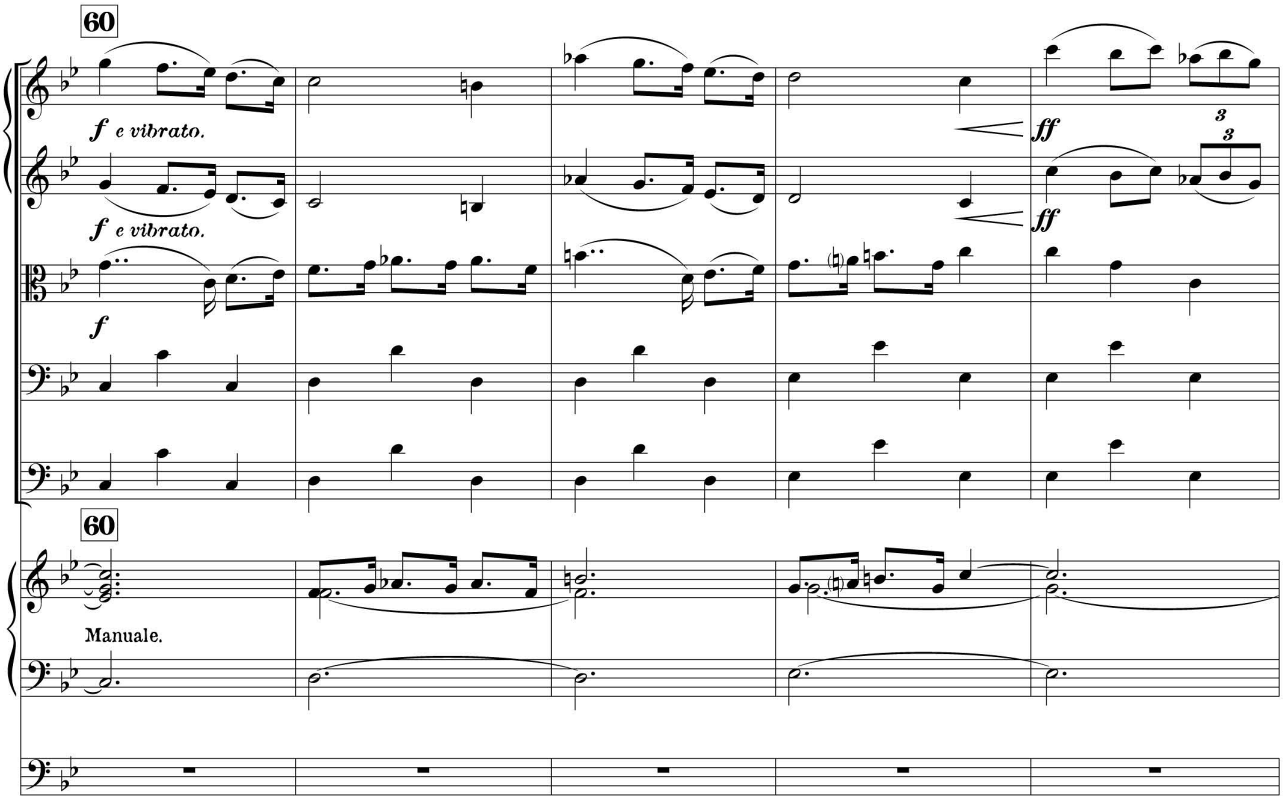
10
p dolce, ma pieno.
pp
p
10
Manuale.
p
mf
15
15

Shorter version's end
(Second refrain)

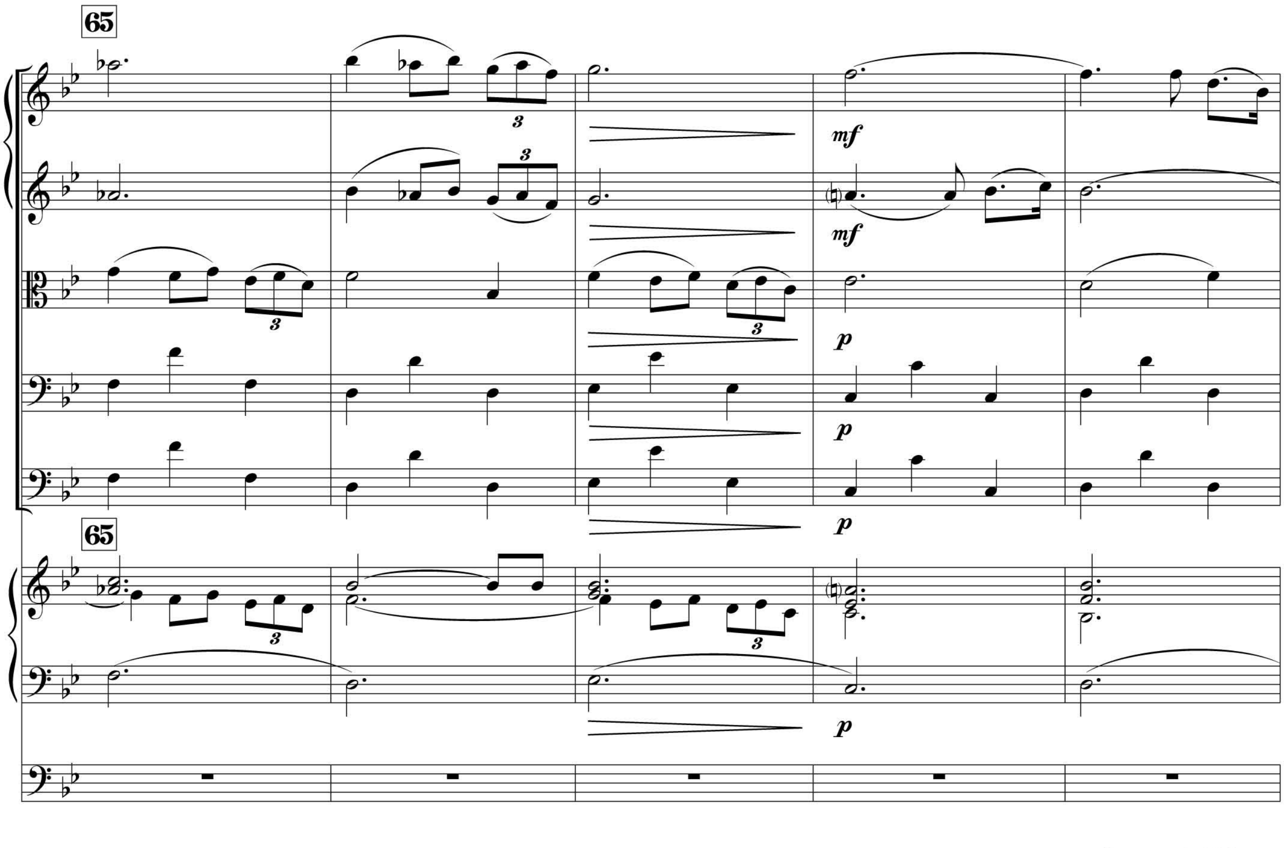


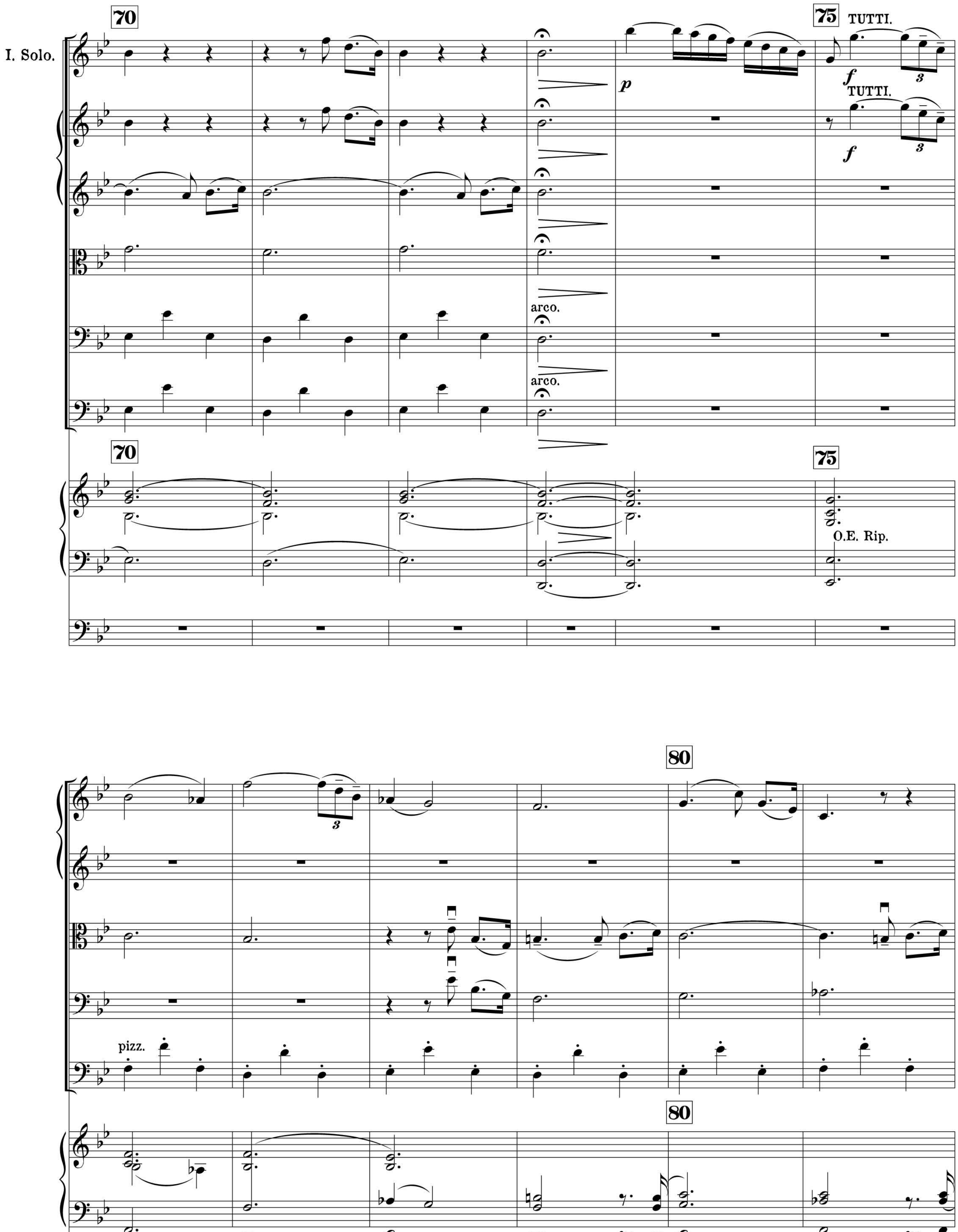

30
I. SOLO (Cadenzando).
I. Solo.
30

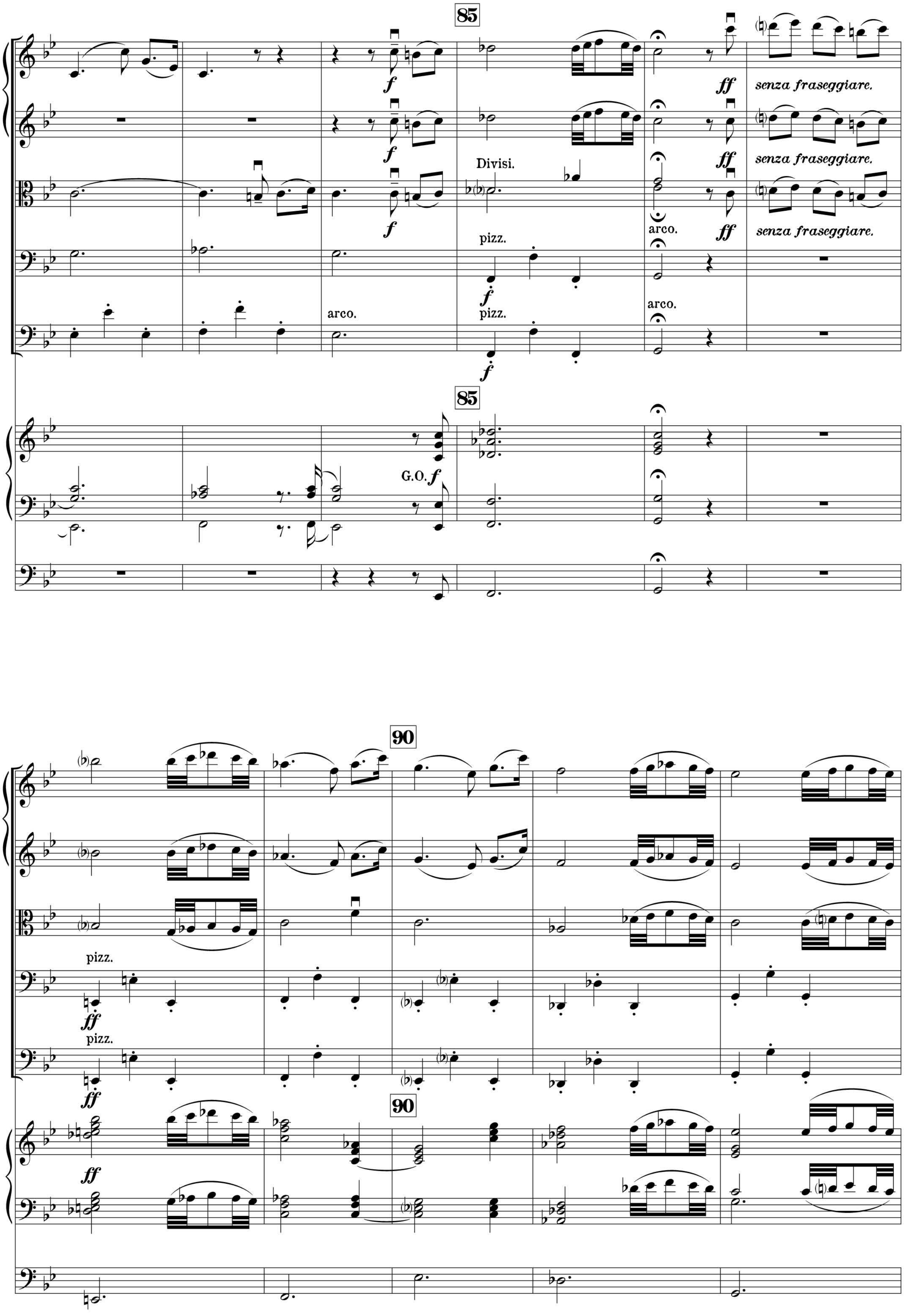

35
I. Solo.
35


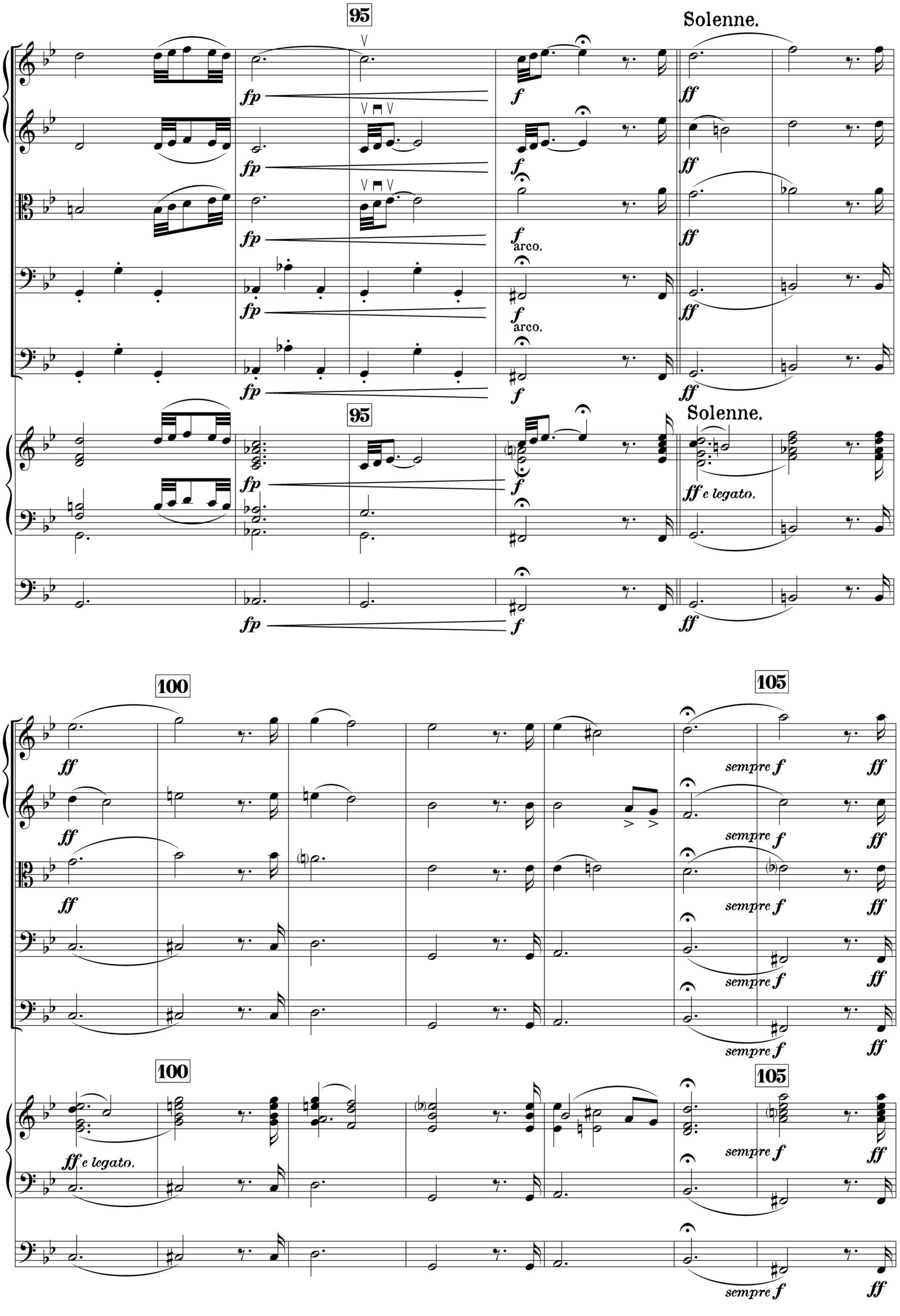
40
I. Solo.
40
45
I. Solo.
45
Manuale.


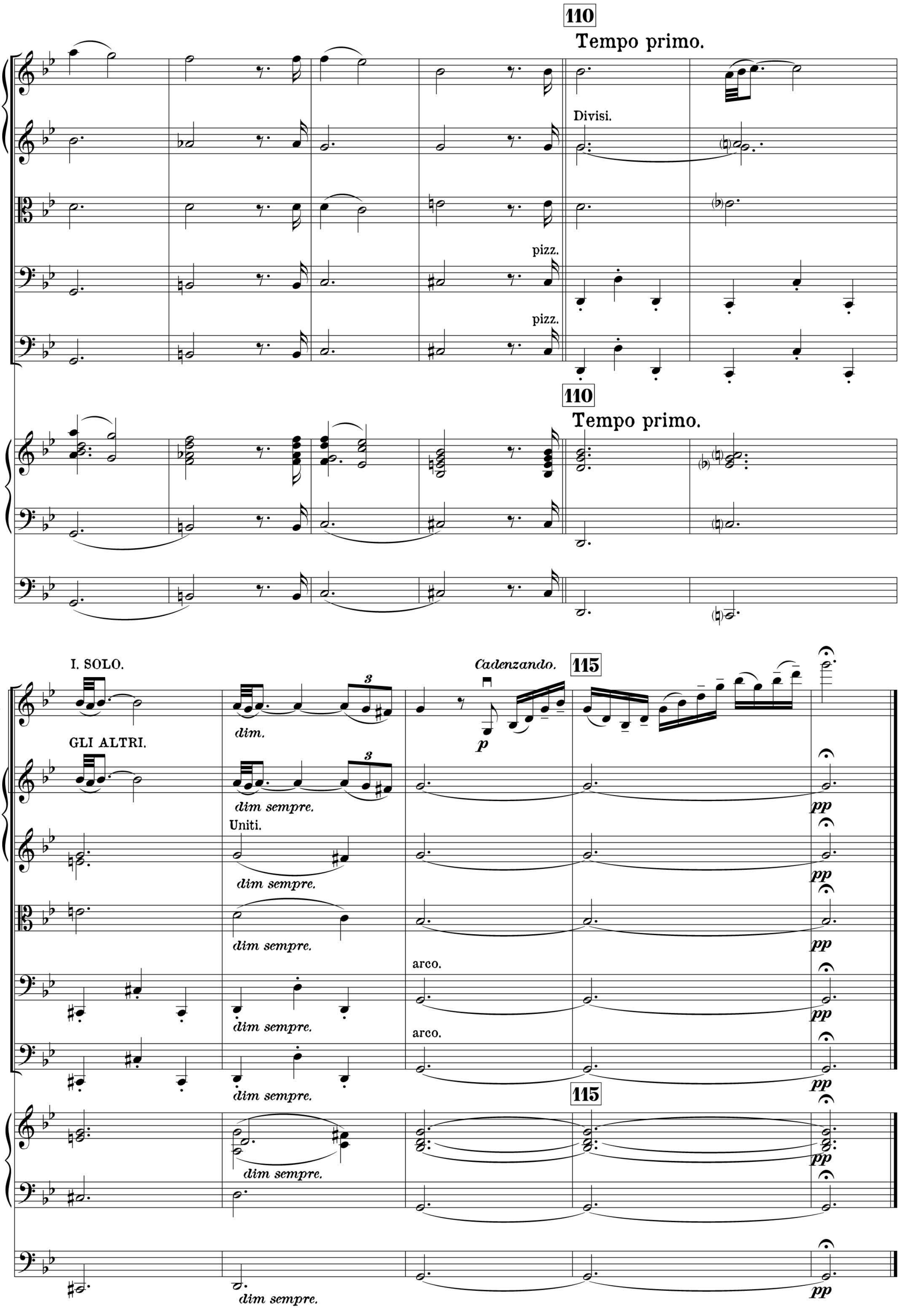
TUTTI.
f
50
p
pp
pizz.
simili.
f
p
pp
pizz.
simili.
f
50
(♮)
f
p
pp
G.O.
55
55
3


60
f e vibrato.
f e vibrato.
f
ff
ff
3
3
60
Manuale.


65
3
3
mf
mf
3
p
p
p
65
3
3
p


70
75
80
I. Solo.
TUTTI.
TUTTI.
p
f
f
3
arco.
arco.
pizz.
O.E. Rip.


85
f
ff
senza fraseggiare.
senza fraseggiare.
senza fraseggiare.
Divisi.
pizz.
arco.
arco.
85
G.O.
f
90
pizz.
pizz.
ff
ff
90


95
Solenne.
fp
f
ff
arco.
ff e legato.
100
105
sempre f


110
Tempo primo.
Divisi.
pizz.
pizz.
110
Tempo primo.
I. SOLO.
I. Solo.
GLI ALTRI.
dim.
dim sempre.
Uniti.
Cadenzando.
115
p
pp
arco.
arco.
dim sempre.
115